Modulation by visual context beyond local features

A DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Cheng Qiu

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisors:
Daniel Kersten
Cheryl A. Olman

December 2015

Acknowledgements

Thanks to my advisors, Daniel Kersten and Cheryl A. Olman. They have enlightened me with their profound knowledge, and their guidance and support have always been with me over the years.

Thanks to my mentors, Sheng He, Stephen Engel, and Gordon Legge, for generously sharing their inputs to both my research projects and academic career. Thanks also go to Galin Jones for his feedback on statistical analysis through serving on my doctoral preliminary and final exam committee.

Thanks to my colleagues and collaborators: Michael-Paul Schallmo, Damien Mannion, Yaniv Morgenstern, Philip C. Burton, Andrea Grant, and Scott Sponheim. Each has taught me great things.

Thanks to Shikha Saggi, Chunye Cen, and Ailin Deng for assistance with research participant recruitment and data collection.

Thanks to my family and friends. Their encouragement and support helped me to go through the hardships in my graduate studies in the US. Most of all, I thank my significant other, Xin Yin, for enriching my life with joy and for assisting to automate tasks in many aspects of my life.

The research was supported by NIH R21NS075525 to Cheryl A. Olman, ONR N000141210883 to Daniel Kersten, P30 NS076408, P41 EB015894 to CMRR, and Eva O. Miller Fellowship, MnDrive Informatics graduate fellowship to Cheng Qiu.

Dedication

To my grandfather, Xiang Zheng.

Abstract

We rarely see an isolated visual stimulus all alone by itself. Rather, the stimulus tends to be surrounded by spatial and temporal context, which often affect both the perception and cortical responses to the target stimulus. The contextual information in fact can be fairly rich, and its effects can be very complex, which in many cases have not been fully explored. This thesis uses functional magnetic resonance imaging (fMRI), psychophysics, and computational modeling to examine the effects and functions of the contextual modulation especially beyond local features.

In Chapters 2 and 3, two sets of experiments considering a larger context including the global shape complexity and figure-ground segregation are used to reconcile competing branches of the literature in terms of the cortical response patterns of early visual areas. In particular, those cortical responses to coherent structure-from-motion stimuli or circular contours vary dramatically depending on the global context. In Chapters 4 and 5, two sets of psychophysical experiments use the tilt illusion and the shape distortion, respectively, as probes to further explore the functions of local and global context.

Overall, the ability to take account larger context ensures the system would dynamically adjust weights between an efficient representation and a strategy to emphasize targets and indicate certainty in early visual areas. Additionally, a stronger perceptual grouping cue between the target and its surround or a larger uncertainty of the center target would enhance the contextual modulation and increase perceptual biases,

which would potentially increase the sensitivity of visual system to feature discrepancies, and it would play an important role in visual search and detection.

Table of Contents

List of Figures

1. Introduction

1. 1 Context in vision

No visual stimulus is in isolation; rather, individual visual components need to be integrated with relevant context and segregated with irrelevance in order to be properly interpreted and used for downstream processes. The surrounding context either in space or in time will then modulate cortical neural responses and visual perception to the central stimuli. Locally, neuronal responses to a stimulus within its receptive field (RF) can be enhanced or inhibited by stimuli outside the classical receptive field (termed as extra-/non-classical receptive field). Various factors such as contrast, orientation, and spatial location would affect this modulation. Further, beyond the non-classical receptive field, even higher-level percepts may also play an important role in modulating local responses, and this is the main focus of the current study. This thesis uses a series of experiments including functional imaging, psychophysics, and computational modeling to probe effects and functions of visual context beyond a local level.

1.1.1 Context discussed locally

Contextual modulation in primary visual cortex (V1) has been intensively studied for decades. Early observations show that the activity of V1 neurons can be modulated by stimuli outside their receptive fields, though these stimuli do not themselves evoke any responses (Blakemore & Tobin, 1972; Cavanaugh, Bair, & Movshon, 2002b; Gilbert & Wiesel, 1990; Knierim & Essen, 1992; C.-Y. Li & Li, 1994; Maffei & Fiorentini, 1976; Nothdurft, Gallant, & Van Essen, 1999; Sillito, Grieve, Jones, Cudeiro, & Davls, 1995).

In the temporal domain, presentation of one stimulus can also modulate the neuronal response to a subsequent test stimulus (Albrecht, Farrar, & Hamilton, 1984; Nelson, 1991). These results are suggested to be physiological explanation for various perceptual effects (Stemmler, Usher, & Niebur, 1995). For example, the perceived orientation could be biased by surrounding stimuli (Blakemore, Carpenter, & Georgeson, 1970; Kapadia, Westheimer, & Gilbert, 2000; O'Toole & Wenderoth, 1977; Wallace, 1969) or previously presented orientations (Gibson & Radner, 1937; Tolhurst & Thompson, 1975; Wenderoth & van der Zwan, 1989). The orientation discrimination threshold would increase due to spatial or temporal context (W. Li, Thier, & Wehrhahn, 2000; Mareschal, Sceniak, & Shapley, 2001; Regan & Beverley, 1985; Wehrhahn, Li, & Westheimer, 1996), whereas the contrast detection threshold may decrease when a pair of collinear flankers is presented (C.-C. Chen & Tyler, 2001, 2002; Polat & Sagi, 1993, 1994; Solomon, Watson, & Morgan, 1999; Woods, Nugent, & Peli, 2002). Direct comparisons between perception and neuronal responses have been made, indicating a good correlation between the two (Kapadia, Ito, Gilbert, & Westheimer, 1995; Kapadia et al., 2000).

Suppressive modulation is thought to regulate the cortical gain and to achieve an efficient representation of visual inputs (Cavanaugh, Bair, & Movshon, 2002a; Heeger, 1992; Schwartz & Simoncelli, 2001; Shapley & Xing, 2013). Additionally, the modulation leads to a larger spatial summation region, allowing selectivity for collinear, cocircular or complex forms such as corners and T junctions (Das & Gilbert, 1999; Field, Hayes, & Hess, 1993; Hess & Field, 1999). The shape of the local contextual modulation areas also reflects the geometric regularity of contours in natural scenes (Geisler, 2008;

Geisler & Perry, 2009; Geisler, Perry, Super, & Gallogly, 2001; Sigman, Cecchi, Gilbert, & Magnasco, 2001), which suggests the essential role of contextual modulations in visual contour integration.

However, the modulations are often more complex than what could be directly modeled using local context. For example, the sign and stimulus selectivity of the modulation from context can be dramatically altered simply by changing the contrast of the center and surround stimuli (Kapadia, Westheimer, & Gilbert, 1999; Levitt & Lund, 1997; Maffei & Fiorentini, 1976; Sceniak, Ringach, Hawken, & Shapley, 1999; Smagt, Wehrhahn, & Albright, 2005; Wehrhahn & Dresp, 1998). Also, diverse visual cues can interactively modulate the neuronal responses even when the cues are not directly relevant. Such flexibility permits dynamic modification of neuronal response properties, which are likely modulated by higher-level process.

1.1.2 Context involved higher-level process

Beyond the local features, high-level perceptual grouping or other features indirectly related to the current feature of interest would modulate neuronal responses and perception through long-range or feedback connections. Three sets of studies are discussed below to highlight the effect of global characteristics on the contextual modulation in V1. First, Lamme (1995) showed that given the identical pattern of textured stimuli, the recorded activities from V1 neurons were determined by figure-ground segregation cues outside their RFs (as shown in Figure 1.1A). Zipser, Lamme, and Schiller (1996) further demonstrated that apart from orientation- and motion-defined figures, the segregation due to other cues, such as binocular disparity, color and

luminance, or even the combination of these cues, evoked very similar modulations in V1 neurons. The invariance in response to diverse cues suggests that the neurons are likely to respond to the higher-level perception of the "figure" generalized across multiple cues, instead of responding to a particular local feature.

The second set of studies indicates that even similar contextual stimuli may not produce similar effects at all, and it highly relies on perceptual grouping or segmentation over the scene. For example, the vernier-offset discrimination is degraded when the vernier stimulus is flanked by two lines or by multiple scrambled line segments, but the reduction can be largely recovered if those segments form a sensible object, e.g. cuboids as shown in Figure 1.1B (Herzog & Fahle, 2002; Herzog & Koch, 2001; Herzog, Thunell, & Ögmen, 2015; Sayim, Westheimer, & Herzog, 2010). Here, if the contextual flankers are grouped with other stimuli rather than the vernier target, the contextual effect would be disrupted. Similarly, Joo and Murray (2014) showed that the orientation sensitivity of the contextual suppression in early visual areas was observed when the center and flankers were grouped on the same surface even with a large distance between them, but not when the center and flankers were assigned to different surfaces (see Figure 1.1C (left), also see Huang, Chen, and Tyler (2012) for an example with collinear facilitation). Interestingly, if the center and flankers jointly form an alternating pattern (Figure 1.1C (right)), the contextual suppression would be very strong (Joo, Boynton, & Murray, 2012). In all the examples, the visual perception and early visual cortical responses to the center target are influenced by a much larger context, and are modulated by perceptual grouping/segmentation.

The perceptual state, visual attention and learning can also modulate the effects of context. Zipser et al. (1996) used dichoptic presentations of the figure-ground textures (Figure 1.1A) and demonstrated that even with a large change in the stimulus, if it had little perceptual effect, it would not alter the neuronal responses in V1 either; that is, the contextual modulation correlated with perceptual experience but not physical visual inputs. Altmann, Bülthoff, and Kourtzi (2003) also showed that the cortical responses in early visual areas could be very similar to the collinear contours (Figure 1.1D (left)) and to the perceptually detected but in fact misaligned contours (Figure 1.1D (right)). Additionally, attention and perceptual learning play an important role in modifying V1 contextual modulation. When the attention is distributed to multiple locations, a larger increase in the response to the target (with flanker) is observed comparing with the condition with the focal attention on the RF; but this effect could be diminished by learning/overtraining (Gilbert, Ito, Kapadia, & Westheimer, 2000; Minami Ito & Gilbert, 1999; Minami Ito, Westheimer, & Gilbert, 1998).

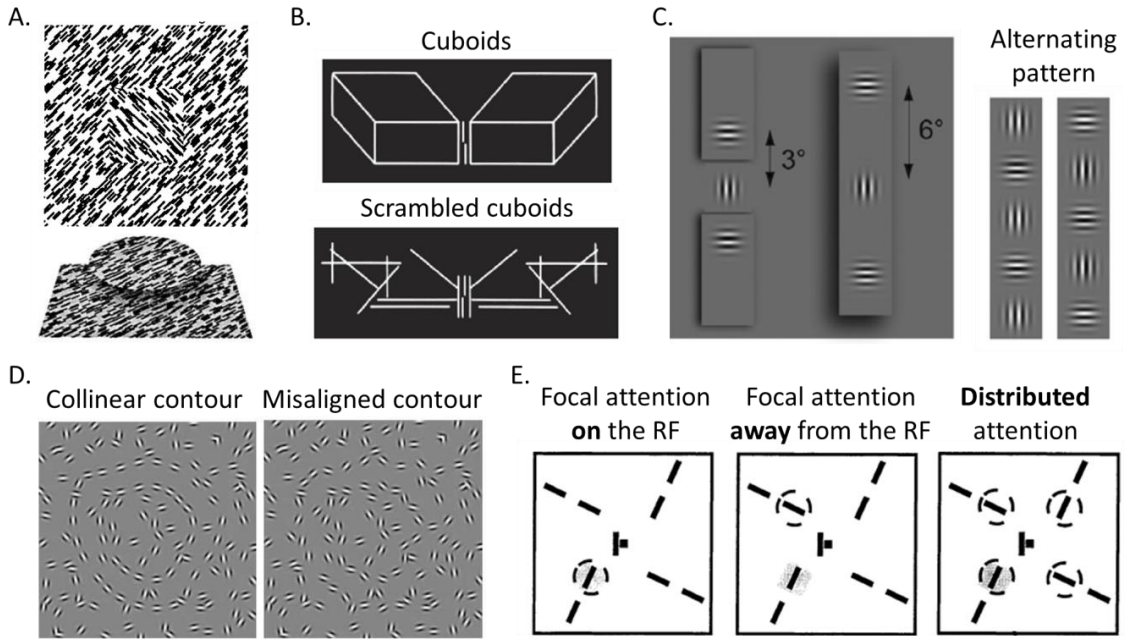

Figure 1.1. Example studies highlighted the effect of global characteristics on the contextual modulation in V1. A. orientation-defined (above) or binocular disparity-defined (below) figure-ground segregation, adapted from Lamme (1995) and Zipser et al. (1996). B. The vernier stimulus surrounded by cuboids or scrambled cuboids, adapted from Sayim et al. (2010). C. Left, the center and flankers on different surfaces or the same surface, adapted from Joo and Murray (2014). Right, an alternating pattern jointly formed by the center and flankers, adapted from Joo et al. (2012). D. Collinear and misaligned contours among background clutter, adapted from Altmann et al. (2003). E. The target line segment with flanker presented with focal attention on the RF, away from the RF, or attention distributed among all targets. The attention is illustrated using dashed circle, and shaded region indicates the RF, adapted from Gilbert et al. (2000).

The sensitivity and flexibility of contextual modulations in early visual areas suggest enormous knowledge of overall scene context in these areas and the ability of

higher-level areas to influence the responses at lower levels. What would be the underlying structure allowing such dedicated process? Visual processing is hierarchical in primates with a sequence of areas in which individual neurons access visual inputs from increasingly large regions and process increasingly complex forms of visual inputs (Felleman & Van Essen, 1991; Markov & Kennedy, 2013; Wallisch & Movshon, 2008). The first stage of visual processing happens in the retina, and the signals are then passed to the lateral geniculate nucleus (LGN), to primary visual cortex (V1) and the second visual area (V2), and then to the extrastriate areas of the dorsal and ventral streams. Areas in the dorsal stream, such as middle temporal complex (MT+) and V3 accessory (V3A), are involved in the processing of visual motion; and ventral areas, such as V4 and lateral occipital complex (LOC), are thought to represent complex visual patterns and objects (DeYoe, Felleman, Van Essen, & McClendon, 1994; Ungerleider & Haxby, 1994). Besides these feedforward connections in cortex, each cortical area receives strong feedback connections from multiple areas: in most cases where there is a forward projection from one area to another, there is a projection in the opposite direction (Felleman & Van Essen, 1991; Markov et al., 2013). The feedback connections provide infrastructure that would permit early visual areas access to perceptual interpretations of the visual scene, such as figure-ground segregation, surface representations and perceptual grouping, from higher-level cortical areas.

Modulations from regions outside the classical RF of neurons either through the intra-areal or inter-areal connections in fact serve as a general rule at various levels in the visual system (Allman, Miezin, & McGuinness, 1985). Exploring the source of these

modulations would allow us a better understanding of the perceptual organization and its processes. In this thesis, focusing on contextual modulations in early visual areas yielded at both local and global levels, we study the role of context in shaping cortical responses and the fundamental goal of the contextual effects on visual perception.

1.2 Major questions and structure of the thesis

With a series of experiments, Murray, Kersten, Olshausen, Schrater, and Woods (2002) demonstrated an opposite response pattern between the lateral occipital complex (LOC) and primary visual cortex (V1) suggesting contextual effects in early visual areas are influenced by grouping processes in higher-level areas. Specifically, cortical response increases in the LOC and concurrent response decreases in V1 are observed when coherent shapes are formed by line segments or by coherently moving dots (also see Hӓndel, Lutzenberger, Thier, and Haarmeier (2007), Dumoulin and Hess (2006), Paradis et al. (2000) and Cardin, Friston, and Zeki (2010) for similar results in V1). However, other studies show response enhancement in V1 to coherent structure relative to scrambled elements (Altmann et al., 2003; Bauer & Heinze, 2002; M. Chen et al., 2014; Kourtzi, Tolias, Altmann, Augath, & Logothetis, 2003; W. Li, Piëch, & Gilbert, 2006; McManus, Li, & Gilbert, 2011; Roelfsema, Lamme, & Spekreijse, 2004); or no sensitivity to structure or motion coherence is observed in V1 (Kriegeskorte et al., 2003; Peuskens et al., 2004). Why does visual context at times suppress, enhance, or have no effect on cortical responses in early visual areas? In Chapter 2 and Chapter 3, two sets of experiments considering context at a more global level are used to reconcile competing branches of the literature in terms of the V1 cortical response patterns to detection of

visual structure. These results reinforce the importance of visual context in modulating responses in early visual areas.

If contextual modulation is a general process at various levels of the visual system, what are the fundamental purposes or what are the basic rules the process needs to follow? In Chapter 4 and Chapter 5, two sets of psychophysical experiments use the tilt illusion and the shape distortion, respectively, as probes to explore the functions of local and global context and to discuss their ecological meaning to visual perception. Overall, such process ensures the system would dynamically adjust weights between an efficient representation in lower-level areas and a strategy to emphasize targets and to indicate certainty. Additionally, a stronger perceptual grouping cue between the center and surround or a larger uncertainty to the center target would enhance the contextual modulation and increase perceptual biases, which would potentially increases the visual system sensitivity to feature discrepancies, or to features otherwise easily confused or degraded. Chapter 6 summarizes the work of this thesis, and future directions are proposed.

2. The effect of global shapes in structure-from-motion perception

2.1 Introduction

Structure-from-motion (SFM) is the perception of three-dimensional structure derived from projected two-dimensional motion of segments (Wallach & O'Connell, 1953). For example, stimuli composed of two-dimensional (2D) moving dots can be perceived as objects rotating in depth (Andersen & Bradley, 1998). Both lesion studies and electrophysiological recordings demonstrated that middle temporal area (MT) plays an essential role in SFM percepts (Bradley, Chang, & Andersen, 1998; Dodd, Krug, Cumming, & Parker, 2001; Newsome & Pare, 1988; Siegel & Andersen, 1990), while functions in the area V1 may be modest (Grunewald, Bradley, & Andersen, 2002). Using functional magnetic resonance imaging (fMRI) in humans, the SFM perception-relevant network has been demonstrated to extend dorsally into the middle temporal complex (MT+) and the parieto-occipital junction (V3A) and ventrally into object-sensitive regions, such as LOC (Kriegeskorte et al., 2003; Orban, Sunaert, Todd, Van Hecke, & Marchal, 1999; Paradis et al., 2000; Peuskens et al., 2004; M. E. Sereno, Trinath, Augath, & Logothetis, 2002; Todd, 2004).

Receptive fields in V1 limit the field of view of a neuron to just a portion of the moving dots (usually smaller than 2 degree in diameter), and hence, it is not possible for one neuron in V1 to signal with certainty the true structure of the object in a purely bottom-up manner. Comparing with area V1, area MT has neurons with large receptive

fields. It is, therefore, capable of spatially integrating motion cues across large angular subtense, about 5-10 degrees (Adelson & Movshon, 1982; Albright & Desimone, 1987; J. A. Movshon, Adelson, Gizzi, & Newsome, 1985; J. A. Movshon & Newsome, 1996); it is also able to distinguish coherent versus incoherent motion (Cheng, Fujita, Kanno, Miura, & Tanaka, 1995; McKeefry, Watson, Frackowiak, Fong, & Zeki, 1997). In addition, area MT is important in tasks of transforming motion cues into information about surfaces and depth (Born & Bradley, 2005; Hildreth, Ando, Andersen, & Treue, 1995; Maunsell & Essen, 1983a, 1983b; McLeod, 1996; Orban, 1997; A. W. Roe, Parker, Born, & DeAngelis, 2007; Snowden, Treue, Erickson, & Andersen, 1991; Treue, Andersen, Ando, & Hildreth, 1995). These characteristics of area MT establish its key role in SFM percepts.

Additionally, areas V3A, V3B, and human V4 (hV4) are recognized as regions at intermediate level in the visual processing hierarchy. Both V3A and V3B are thought to be involved in processing motion information (A. T. Smith, Greenlee, Singh, Kraemer, & Hennig, 1998; Tootell et al., 1997); whereas hV4 in the ventral stream is sensitive to complex visual forms, such as high curvature and partially occluded forms (Bushnell, Harding, Kosai, & Pasupathy, 2011; Nandy, Sharpee, Reynolds, & Mitchell, 2013; Pasupathy & Connor, 2002). Further, area LOC is specialized in cue-invariant structure perception (Fang, Kersten, & Murray, 2008; Grill-Spector, Kourtzi, & Kanwisher, 2001; Grill-Spector, Kushnir, Edelman, Itzchak, & Malach, 1998; Grill-Spector, Kushnir, Hendler, et al., 1998; Gross, Rocha-Miranda, & Bender, 1972; Haxby et al., 2001;

Kourtzi & Kanwisher, 2000, 2001; Logothetis & Sheinberg, 1996; Moore & Engel, 2001; Tanaka, 1996), which is an important component in SFM percepts as well.

Response increases to coherent versus scrambled structure in higher-level areas are almost certain among existing literature, but response patterns in early visual areas often vary in different studies. For example, with spectral analysis in a MEG study, the component attributed to areas including MT increased linearly with motion coherence; while another component from early visual cortex showed the inverse dependence on motion coherence (Händel et al., 2007). Similarly, Paradis et al. (2000) and Murray et al. (2002) observed response reduction in area V1, but response increase in area LOC using fMRI when coherent SFM stimuli were presented. However, other results showed that early visual areas responded approximately equally to moving dots (Kriegeskorte et al., 2003).

In order to reconcile these results, two levels of contextual modulations were considered in the current study. At the first level, the motion coherence was controlled, so either coherent structures or scrambled-dot motion would be perceived. At the second level, the familiarity of the perceived shapes was manipulated. In both cases, local stimulus features such as motion vectors and texture gradients of the moving dots were largely maintained to ensure that area V1 receives similar ascending signals, and thus modulations in V1 activity were most likely attributed to long-range horizontal connections or inter-areal feedback.

Two functional imaging experiments were conducted using SFM stimuli with either unfamiliar/complex (Experiment 1) or simple (Experiment 2) structures. Two

conditions were included within each experiment: in the coherent condition, transparent three-dimensional (3D) structures rotating in depth could be perceived; in the scrambled condition, motion coherence and opponent 2D motion are maintained but no 3D structure can be perceived. In both experiments, area V3A/B and LOC showed significant preference to the coherent structure; however, the response reduction in V1 was only observed in Experiment 2 when the structure is simple. These results suggest that both local and global context would influence cortical responses in early visual areas.

2.2 Methods

2.2.1 Participants

Eight observers (mean age: 30 years old, four males) and five observers (mean age: 28 years old, three males) with normal or corrected-to-normal visual acuity participated in Experiment 1 and Experiment 2, respectively. The observers provided informed written consent under an experimental protocol that was in accordance with safety guidelines for MRI research and was approved by the Institutional Review Board at the University of Minnesota.

2.2.2 Visual stimuli

Experiment 1. Novel 3D structures were generated using perturbations in spherical harmonics (Figure 2.1A). The 3D positions of the dots were randomly selected over the structure surface, and were presented under parallel projection. The projected dots would sometimes be adjusted to ensure that the dot density was approximately uniform (Figure 2.1B). The dot trajectories were computed by projection of this set of dot positions as the structure moving in 3D. The structure was modeled as transparent with the dots

continuously visible even when they moved to the back surface. For each presentation (900 ms), the structure would rotate about a random axis in 3D space for about 40 °, and the starting frame was also random.

Experiment 2. A simple structure, cube, was used to substitute the novel shapes in Experiment 1, and other parameters were maintained.

In both experiments, displays consisted of ~450 moving dots occupied about 10 ° visual filed. The dots were round with a radius of 0.08 °visual angle. Each stimulus contained a fixation region with a diameter of 1 °visual angle at the center of the screen. Two conditions, scrambled and coherent SFM, were included. Low-level stimulus features were identical for both. In the coherent SFM condition, parameters that induced stable SFM percepts based on Siegel and Andersen (1988) were applied. In the scrambled conditions, the motion vector for each dot was maintained, but the position of the dot was randomly shifted with a small amount. As a result, the perception of a coherent surface was eliminated, and, instead, a cloud of dots would be seen rotating in depth without a 3D surface or structure. In this way, local motion coherence (percentage of dots locally moving in the same direction) and opponent motion (opposite local motion direction) were well maintained. Within a small area of the stimuli, the coherent and the scrambled conditions show almost no difference. However, the perceptions of these stimuli were quite different. Figure 2.1C&D illustrate the stimuli in both experiments for scrambled and coherent conditions using a summation of three continuous frames.

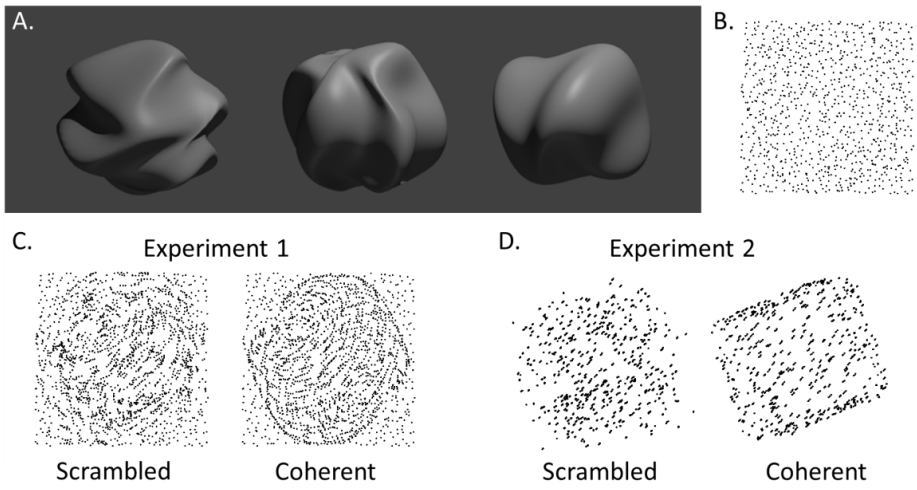

Figure 2.1. Examples of stimuli used in the structure-from-motion experiments. A. Novel 3D shapes used in Experiment 1. B. One frame of the actual stimulus in Experiment 1, where low-level stimulus features were tightly controlled. C. Illustration of stimuli in Experiment 1 with a novel 3D shape. Sum of 3 continuous frames are shown to illustrate conditions at varying coherence levels. D. Illustration of stimuli in Experiment 2 with a cube.

2.2.3 fMRI experiments

Stimuli were displayed using a Sony video projector (spatial resolution of 1024 × 768 pixels, temporal resolution of 60 Hz, and mean luminance of 168 cd/m^2) on a translucent screen positioned within the scanner bore. Observers viewed the stimuli from a distance of 72 cm through a mirror located above their eyes (mounted on the head coil), which gave a total image area subtending 29.1 ° × 21.8 °. Stimuli were presented using Matlab (R2010b; Mathworks, Inc., Natick, MA, USA) with the Psychtoolbox extensions (Brainard, 1997; Pelli, 1997). Behavioral responses were collected via a FIU-005 fiber optic response device (Current Designs, PA).

Functional MRI data were collected using a 7T scanner (Siemens, Erlangen, Germany) with a head gradient set. A T_2^* sensitive gradient echo imaging pulse sequence was used: TR = 2 s, TE = 18 ms, flip angle = 70 °, matrix = 108 x 108, GRAPPA acceleration factor = 2, FOV = 162 x 162 mm, partial Fourier = 7/8, voxel size = 1.5 mm isotropic, and 36 coronal slices were obtained to cover the occipital lobes (Figure 2.2A).

A scanning session contained four runs of block-design experiments. Three conditions including scrambled/coherent SFM stimuli and a control condition with static frames were randomly interleaved. Each block lasted 16 sec, in which 16 different rotations were presented each with 1 sec (100 ms static at the end for each rotation). In the control condition, 16 random static frames were shown for each block. Observers were instructed to perform a mental calculation task at the fixation region.

2.2.4 Anatomical acquisition and visual area mapping

A T_1-weighted anatomical image (sagittal MP-RAGE, 1 mm isotropic resolution) was acquired for each participant in a separate session using a Siemens Trio 3T magnet. FreeSurfer (Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999) was used for segmentation, cortical surface reconstruction, and surface inflation and flattening of the anatomical image. The warped surface was then converted to a standard mesh using SUMA (Saad, Reynolds, Argall, Japee, & Cox, 2004). Visual areas were defined in a separate scanning session following standard procedures, including four runs of a clockwise/counter-clockwise rotating wedge stimulus plus two runs of an expending/contracting ring stimulus to identify the visual areas V1, V2, V3, V3A/B, and hV4 (DeYoe et al., 1996; Engel, Glover, & Wandell, 1997; Larsson & Heeger, 2006;

Wade, Brewer, Rieger, & Wandell, 2002), two runs contrasted blocks of translating and static low-contrast dots to locate the MT+ (Tootell et al., 1995; Zeki et al., 1991), and two runs contrasted scrambled and coherent objects for LOC (Kourtzi & Kanwisher, 2001; Malach et al., 1995). Figure 2.2B shows an example angular map with identified visual areas labeled. Area V3B appeared to share the central representation with V3A (Press, Brewer, Dougherty, Wade, & Wandell, 2001; Wandell, Brewer, & Dougherty, 2005), and the boundary between the two was not clear, we would combine them as V3A/B in the analysis below.

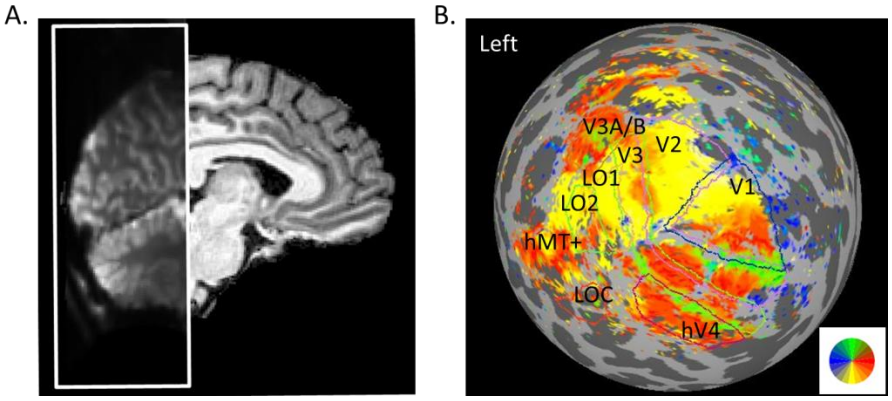

Figure 2.2. Example functional image and retinotopic map. A. Sagittal view of example functional image coverage. The image within the white box shows a session average of pre-processed functional data. B. Example angular visual field map from a left hemisphere (inflated cortical surface visualized as a sphere) with identified visual areas labeled.

2.2.5 Pre-processing

Motion correction parameters acquired using AFNI (Cox, 1996) with reference to the volume right before a within-session fieldmap image were first combined with distortion unwarping parameters via FSL (S. M. Smith et al., 2004). Functional imaging

data were then resampled with sinc interpolation. Individual observer's anatomical image was aligned with a mean of all the functional images with AFNI's align_epi_anat.py using six free parameters for translation and rotation. Please see Mannion, Kersten, and Olman (2014) for more details. The pre-processed functional data were then projected onto the standardized cortical surface. All analysis was performed on the nodes of this surface domain representation.

2.2.7 Analysis

A general linear model (GLM) was used to analyze the functional data. Stimulus blocks for each condition were modeled as boxcars and convolved with SPM's canonical hemodynamic response function. The GLM was estimated using AFNI's 3dREMLfit, which estimates and removes noise temporal correlations with a voxelwise ARMA(1,1) model. The stimulus condition beta estimates for each node from the GLM were converted to percent signal change via division by the estimate baseline timecourse from nuisance regressors. These percent signal change values were than averaged across the nodes that significantly responded to the moving conditions within each predefined visual area. Results were normalized to the control (static) condition for each experiment. Statistical significance for BOLD response differences between conditions was assessed using a bootstrap procedure in which observers' data were resampled with replacement across 10000 iterations. The *p*-values (two-sided test) were corrected using Bonferroni procedure.

2.3 Results

The surface nodes responsive to the moving conditions were first identified (within the ~10 °central visual filed). In detail, for each observer, the nodes with a significantly positive ($p < 0.001$, hemisphere FDR corrected) coefficient for the contrast between the sum of the response to moving stimuli and the response to the static control stimuli were selected. The responses to the coherent and scrambled SFM stimuli were then averaged within selected nodes of each predefined visual area. The BOLD response differences between the two conditions are shown in Figure 2.3 for both experiments. A positive number indicates a stronger response to the coherent condition than the scrambled, and a negative number shows a preference to the scrambled condition.

Among the early visual areas, V1 and V2 showed significant modulations only in Experiment 2 with the Bonferroni adjusted $p < 10^{-16}$ (bootstrap two-sided test) for both areas, and the mean response difference was -1.65% in V1 and -1.15% in V2. In contrast, the uncorrected *p*-values in area V1 and V2 in Experiment 1 were 0.176 and 0.036, respectively. For dorsal regions, V3AB showed a significant response increase to the coherent SFM condition in both experiments with corrected $p < 10^{-16}$, and the response difference was 0.43% in Experiment 1 and 0.57% in Experiment 2. MT+ had a weak preference to the coherent condition in Experiment 2 (corrected $p = 0.040$ with mean response difference 0.42%). In ventral areas, hV4 revealed slight but consistent response increase (mean as 0.24%, corrected $p = 0.0028$) to the coherent condition in Experiment 1; LOC showed strong preference to the coherent condition in both experiments ($p < 10^{-16}$, and mean as 0.72% and 0.78%, respectively).

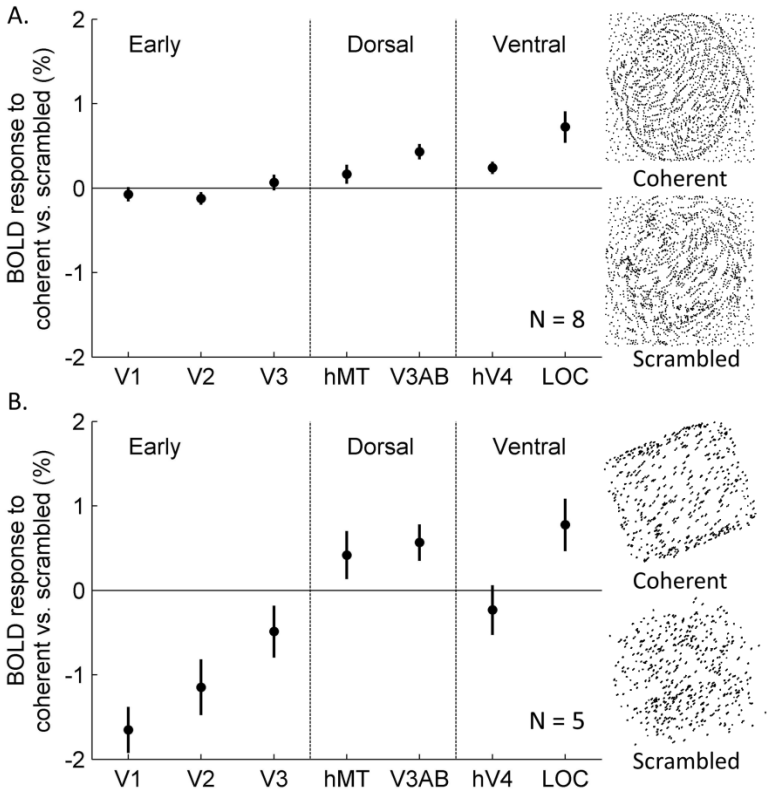

Figure 2.3. Results of the BOLD response contrast between coherent and scrambled SFM stimuli. A. Results from Experiment 1 when the structures are novel 3D shapes generated using perturbations in spherical harmonics. B. Results from Experiment 2 when the structure is a cube. Error bars show ±1 *SE*.

2.4 Discussion

We investigated how global structure context would affect the responses in human visual cortex (early visual areas and some dorsal/ventral areas) using fMRI. Our results show that the effect of motion coherence would also be modulated by the global shape. In Experiment 2, we replicated the response reduction in early visual areas (V1 and V2) when a coherent simple cube was perceived; however, no similar effect was observed in Experiment 1 when the novel 3D structures were presented.

Murray et al. (2002) proposed a possible explanation for the response reduction in V1 with coherent perception: the predictive coding theory (Mumford, 1992; Rao & Ballard, 1999) would predict such reduction due to the decrease of unexplained residuals in V1 when visual inputs were successfully modeled as coming from a coherent shape. Our results would argue that the reduction is selective: given an easy target, though the constant lower-level inputs are suppressed, the structure still could be accurately represented in the high-level areas; whereas when the perceived shape was complicated, though the structure had already been sensed by the higher-level areas (such as LOC), constant lower-level inputs were still necessary for the system to encode the novel shape.

Area hV4 also appeared to be sensitive to the shape familiarity, in which a significant preference to the coherent structure was observed in Experiment 1 but not in Experiment 2. More resources were engaged when the visual structure was complex or unfamiliar. In contrast, area MT+ was less responding to the shape information: it equally responded to both moving conditions in Experiment 1, and only showed a weak trend in preferring coherent structure in Experiment 2.

Additionally, we found that V3A/B may be at a higher hierarchical level than MT+ in the dorsal stream because of a more consistent preference for coherent SFM in V3A/B than MT+, and this hierarchy is also anatomically predicted by Markov et al. (2013). A similar relationship has been demonstrated between ventral areas hV4 and LOC: hV4 is selective to intermediate-level features, whereas LOC responds to global shapes (Kourtzi & Kanwisher, 2001; Lerner, Hendler, Ben-Bashat, Harel, & Malach, 2001; Pasupathy & Connor, 2002). In our results, both V3A/B and LOC had more

consistent preference to coherent moving structures independent on familiarity, suggesting their crucial role in global shape representations. This result was in fact consistent with many studies that have reported greater sensitivity to 3D surface in V3A (Grill-Spector et al., 2001; Grill-Spector, Kushnir, Edelman, et al., 1998; Grill-Spector, Kushnir, Hendler, et al., 1998; Paradis et al., 2000; M. E. Sereno et al., 2002).

To summarize, we have shown that the responses of human early visual areas to structures relying on both their coherence levels and global shape familiarity. This ensures the system flexibility to switch or dynamically weight between strategies to both efficiently and accurately represent visual inputs.

3. Responses in early visual areas to contour integration are context dependent

Authors

Cheng Qiu, Philip C. Burton, Daniel Kersten, Cheryl A. Olman

Summary

It has been shown that early visual areas are involved in contour processing. However, it is not clear how local and global context interact to influence responses in those areas, nor has the inter-area coordination that yields coherent structural percepts been fully studied, especially in human observers. In this study, we used functional Magnetic Resonance Imaging (fMRI) to measure activity in early visual cortex while subjects performed a contour detection task in which alignment of Gabor elements and background clutter were manipulated. Six regions of interest (two ROIs, containing either the cortex representing the target or the background clutter, in each of areas V1, V2, and V3) were predefined using separate target versus background functional localizer scans. The first analysis using a general linear model (GLM) showed that when background clutter was absent, responses in V1 and V2 target ROIs were suppressed by aligned contours compared with unaligned, whereas in the presence of background clutter, responses were significantly stronger to aligned than unaligned contours. The second analysis using inter-area correlations showed that with background clutter, there was an increase in V1-V2 coordination within the target regions when perceiving aligned versus unaligned contours; without clutter, however, correlations between V1 and V2 were

similar no matter whether aligned contours were present or not. Both the average response magnitude and the connectivity analysis suggest different mechanisms support contour processing with or without background distractors. Coordination between V1 and V2 may play a major role in coherent structure perception, especially with complex scene organization.

Keywords: contour integration, fMRI, early visual areas

3.1 Introduction

Contour integration involves grouping local features across several levels of abstraction and a range of spatial scales. Small, similar elements positioned closely along an invisible smooth path are perceptually organized as due to a continuous contour. This grouping process is enhanced if the elements have orientations that align with the path (Field et al., 1993; Hess & Field, 1999; Kovács, 1996; W. Li & Gilbert, 2002). Further, knowledge of the global form of the path contributes to local integration, such as the form closing (Kovács, 1996; Kovács & Julesz, 1993, 1994) and smoothness (Pettet, 1999; Pettet, McKee, & Grzywacz, 1998), and the global knowledge is often necessary to disambiguate competing local groupings in cluttered scenes (Ullman & Sha'ashua, 1988).

It has been repeatedly reported that cortical areas higher in the visual hierarchy such as the lateral occipital complex (LOC) and inferotemporal cortex (IT) show selectivity to coherent contours (Altmann et al., 2003; Cardin et al., 2010; Dumoulin, Dakin, & Hess, 2008; Dumoulin & Hess, 2006; Kourtzi et al., 2003; Mendola, Dale, Fischl, Liu, & Tootell, 1999; Murray et al., 2002; Tanskanen, Saarinen, Parkkonen, & Hari, 2008), which is also consistent with results of shape or object perception from those areas (Fang et al., 2008; Grill-Spector et al., 2001; Grill-Spector, Kushnir, Edelman, et

al., 1998; Grill-Spector, Kushnir, Hendler, et al., 1998; Gross et al., 1972; Haxby et al., 2001; Kourtzi & Kanwisher, 2000; Logothetis & Sheinberg, 1996; Tanaka, 1996). However, responses to similar contour stimuli in early visual areas are still controversial. While most neurons in the primary visual cortex (V1) are thought to have small receptive fields, it is known that their responses are modulated by contextual information from outside their receptive fields (Allman et al., 1985; Fitzpatrick, 2000). However, there are substantial conflicting results on how scene context or global perception affects responses in early visual areas, including V1.

Both enhancement (Altmann et al., 2003; Bauer & Heinze, 2002; M. Chen et al., 2014; Kapadia et al., 1995; Kourtzi et al., 2003; W. Li et al., 2006; McManus et al., 2011; Roelfsema et al., 2004) and suppression (Cardin et al., 2010; Dumoulin & Hess, 2006; Murray et al., 2002; Murray, Schrater, & Kersten, 2004) of cortical responses to coherent contours relative to scrambled elements have been observed in V1 and/or other early visual areas. Recording from individual V1 neurons in monkeys, Kapadia et al. (1995) and W. Li et al. (2006) showed that multiple randomly placed and oriented line segments outside the neuron's receptive field would inhibit its response to an optimally oriented line within its receptive field; however, once some of the surround segments were placed collinearly with the central line, the response was then facilitated. Similar results have been demonstrated using functional neuroimaging in both monkeys and human subjects (Altmann et al., 2003; Kourtzi et al., 2003): cortical responses from early visual areas including areas V1, V2, and V3 showed selectivity to coherent patterns of closed contours embedded in a field of randomly oriented segments.

In contrast, Cardin et al. (2010) showed lower activity in areas V1/V2 for collinear patterns than for noncollinear ones. Dumoulin and Hess (2006) also showed weaker activity in early visual areas to a 100% coherence circular pattern but stronger activity to scrambled patterns. In a third study, Murray et al. (2002) used line-drawings as stimuli and showed smaller responses in V1 to lines that formed two-/three-dimensional shapes than random lines. These were not the only results showing deactivation in early visual areas to stimulus regularities – early visual areas also respond less to coherent than incoherent motion (Händel et al., 2007; Harrison, Stephan, Rees, & Friston, 2007; McKeefry et al., 1997). A similar trend has been observed with coherent versus scrambled natural images using functional imaging in human observers (Grill-Spector, Kushnir, Hendler, et al., 1998; Lerner et al., 2001; Paradis et al., 2000).

We are interested in how the enhancement and suppression of cortical responses in early visual areas to coherent contours could both be true. In either case, neurons in V1 are responding to the same coherent circular contour, but they show different response patterns based on different studies listed above. We think the response pattern may highly depend on stimulus context: results showing response increase mainly used contours embedded in clutter (i.e., with randomly orientated and placed segments in the background), but the ones showing a decrease were usually using isolated structure or if any, with just uniform background. To test this, we designed a two-by-two experiment where context and contour coherence were manipulated, and we used fMRI to record the Blood Oxygenation Level-Dependent (BOLD) signal from the retinotopically corresponding regions in early visual areas (V1, V2, and V3). Subjects performed a

contour detection task in the experiment. When there was no background clutter present, the BOLD responses in regions corresponding to the location of the target were suppressed by aligned contours compared with the unaligned; however, with background clutter, the response trend was reversed. By analyzing the effect of context and contour coherence on correlations between responses in early visual areas, we further demonstrated that the coordination between the retinotopically relevant regions in V1 and V2 was dependent upon the experimental conditions. These results suggest the involvement of multiple strategies in contour integration, and they interact according to the context.

3.2 Materials and methods

3.2.1 Participants

Twelve observers (mean age: 29 years old, seven males) with normal or corrected-to-normal visual acuity participated in the study. The observers provided informed written consent under an experimental protocol that was in accordance with safety guidelines for MRI research and was approved by the Institutional Review Board at the University of Minnesota.

3.2.2 Stimuli

Figure 3.1 provides examples of the stimuli used, which were generated and presented with Matlab (R2010b; Mathworks, Inc., Natick, MA, USA) using the Psychtoolbox extensions (Brainard, 1997; Pelli, 1997). The target region consisted of 8 Gabor patches, which were centered at equal intervals along the circumference of an invisible circle centered at fixation with a radius of 2 °, that is, these patches were evenly

spaced at 2 °eccentricity from the fixation. Each Gabor patch consisted of a 4 cycles-per-degree (cpd) sinusoidal grating with a random phase offset modulated by a Gaussian envelope with full width at half-maximum of 0.4 °(σ=0.17 °). In an aligned condition, the grating orientation of each Gabor patch was aligned with the tangent line to the invisible circle at the point of patch center. These Gabor patches therefore could be perceived as forming a complete circle. In an unaligned condition, the grating orientation of each Gabor patch was randomly generated. The background clutter consisting of the same Gabor elements as the target was located along invisible circles centered at fixation and with radii of 1.2 °, 2.9 °and 4.0 °. Along each circle, the distance between every two Gabor patches was the same as in the target region (1.6 °), thus the numbers of Gabor patches along each eccentricity were 5, 12 and 16, respectively. The grating orientations of these background patches were randomly generated. All Gabor patches had an 80% Michelson contrast and were presented on a mean gray background. The four experimental conditions included aligned target only, unaligned target only, aligned target with background and unaligned target with background.

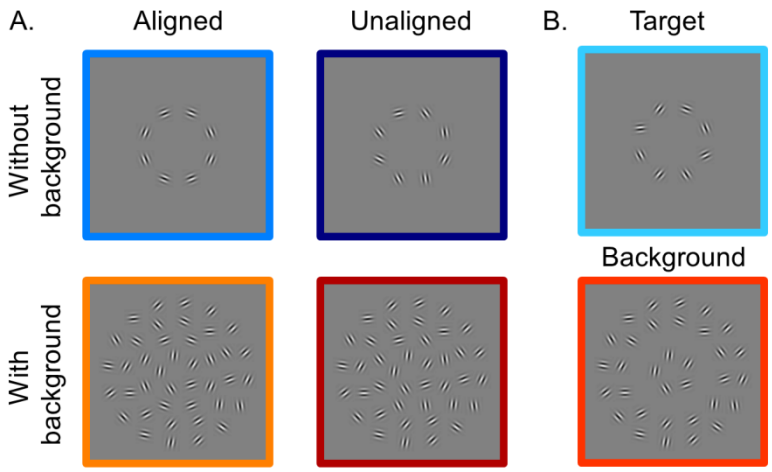

Figure 3.1 Example stimuli. A. Example stimuli from the four test conditions with a circular contour detection task. B. Example stimuli from the target (above) versus background (below) differential localizer scan.

3.2.3 fMRI experiments

Stimuli for the first four subjects were presented on a NEC 2190UXi monitor with resolution of 1024 × 768 pixels and a refresh rate of 60 Hz. The monitor had a mean luminance of 110 cd/m^2. The monitor was mounted to the back wall of the scanning suite; observers viewed the monitor through a mirror mounted to the top of the head coil so that it subtended 12 °of visual angle in the horizontal direction and 9 °in the vertical direction. For the rest of the observers, stimuli were back-projected via a Sony video projector (spatial resolution of 1024 × 768 pixels, 60 Hz refresh rate and 120 cd/m^2 mean luminance) on to a translucent screen placed inside the scanner bore. Observers viewed the stimuli from a distance of 97.5 cm through a mirror located above their eyes (mounted on the head coil), which gave a total image area subtending 26 ° × 20 °.

Functional MRI data were collected using a 3T Siemens Trio scanner (Erlangen, Germany) with a 12-channel head array coil. EPI data were acquired with a field of view 128 mm ×256 mm and a matrix size of 64 ×128 for an in-plane resolution of 2 mm ×2 mm. Slice thickness was 2 mm without inter-slice gap, and number of slices was 20. Echo time (TE) was 30 ms, repetition time (TR) was 1.5 s, and flip angle was 80 °. Four out of the twelve datasets were collected in an axial direction, and the other eight were in a coronal orientation. Both covered the early visual areas V1, V2, and V3.

A scanning session (Figure 3.2A) contained three runs (run 1, 5 and 8) of block-design functional localizers and five event-related runs (run 2, 3, 4, 6 and 7). Observers were instructed to maintain their fixation on a white square at the center while performing behavioral tasks during both localizer and event-related scans. Behavioral responses were recorded using a fiber-optic button box (Current Designs, Philadelphia, PA, USA).

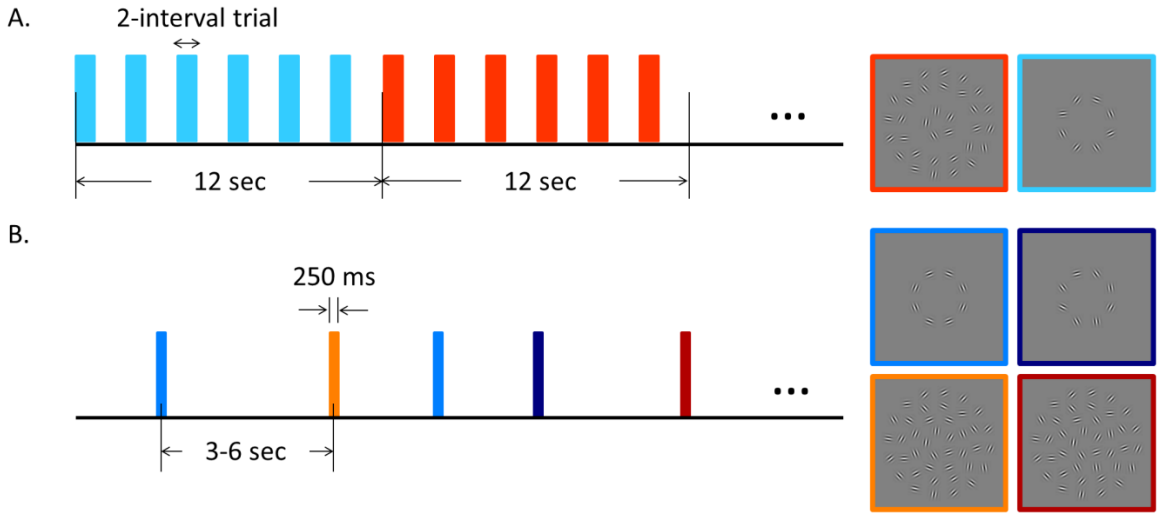

Figure 3.2. Experimental procedure. A. An example of block-designed functional localizer scans used to define target or background retinotopically corresponding ROIs. B. Event-related scans with four experimental conditions.

Regions of interest (ROIs) were defined by block-designed functional localizer scans (Figure 3.2A). Each block lasted 12 s, and each run contained 11 “on” blocks alternating with 10 “off” blocks. Each localizer scan thus lasted 252 s. During “on” blocks randomly oriented Gabor patches at the target region were presented; during “off” blocks only clutter Gabor patches at the background region were presented. The first half-cycle “on” block was discarded before analysis, and the remaining blocks were alternating in 10 cycles per run. During each block, a two-interval trial occurred every 2 s (six trials per block). Duration for both intervals was 200 ms, which were separated by a 200 ms inter-stimulus interval (ISI). Observers were instructed to press button when the stimuli from two intervals were the same (with a probability of 12.5%). The fixation square turned green to a correct response and red otherwise.

The event-related runs (Figure 3.2B) measured BOLD response to four experimental conditions: aligned target only (alnb, aligned/no background), unaligned target only (uanb, unaligned/no background), aligned target with background (albg) and unaligned target with background (uabg). Stimulus duration was 250 ms, and inter-trial intervals (ITIs) were 3, 4.5, or 6 s. ITI was randomly assigned and uniformly distributed. Each run included 20 trials for each condition, thus a total of 80 trials per run. The average run length was about 380 s. Observers were required to press the blue button if they perceived an aligned circle, and press the red button if not. Feedback was provided after each trial.

3.2.4 Anatomical acquisition and Visual area mapping

Prior to the fMRI experiments each observer participated in a separate retinotopic mapping session, in which a T_1-weighted anatomical image (MP-RAGE, 1 mm isotropic resolution) was also collected for anatomical reference and cortical surface definition. Gray/white matter segmentation, cortical surface reconstruction, and surface inflation and flattening were completed using FreeSurfer (Dale et al., 1999; Fischl et al., 1999). Standard retinotopic mapping including four runs of clockwise/counterclockwise rotating wedges and two runs of expanding/contracting rings (DeYoe et al., 1996; Engel et al., 1997; M. I. Sereno et al., 1995) was used to identify the early visual areas V1, V2, and V3. Defined visual areas were registered to the reference anatomy for each observer.

3.2.5 Pre-processing and functional localizers

Functional data was motion corrected using Analysis of Functional NeuroImages software (AFNI) (Cox, 1996), with reference to the volume right before a within-session fieldmap image. The motion-corrected data was unwarped using FSL FUGUE to correct distortions introduced by magnetic field inhomogeneities (S. M. Smith et al., 2004). High-pass filtering was also applied to the functional localizers data: temporal frequencies below four cycles per run were removed. The pre-processed functional data were then aligned to anatomical reference data using mrAlign implemented in Matlab (http://gru.brain.riken.jp/doku.php/mrTools/overview).

ROIs were defined based on both retinotopic visual areas and functional localizers for each observer. Three repetitions of the functional localizers were averaged to define the ROIs responding to the target (tg) or background (bkgd) regions. For each voxel coherence (unsigned correlation, computed in the Fourier domain as the amplitude of the

stimulus-related Fourier component normalized by the square root of the integrated power spectrum) with a sinusoid at the block-alternation frequency, 10 cycles per run, was calculated in the averaged localizer scans (Bandettini, Jesmanowicz, Wong, & Hyde, 1993; Engel et al., 1997). The voxels with coherence exceeding 0.30 were included in the ROIs. The voxels in phase with the target representation were assigned to target ROIs (tgROIs), and the voxels in phase with the background representation were assigned to background ROIs (bkgdROIs). ROIs were initially defined on a flattened cortical surface, where V1, V2 and V3 boundaries could be used to identify the ROIs in different visual areas. Selected voxels were translated to the in-plane space for further refinement to include only contiguous clusters of visually responsive voxels, and the defined ROIs were then exported as a binary mask in the space of the functional data. Six ROIs were defined for each observer: tgV1, tgV2, tgV3, bkgdV1, bkgdV2, and bkgdV3.

3.2.6 Analysis of the event-related data

fMRI data analysis

Functional image analysis of the event-related runs was conducted using general linear model (GLM) in AFNI with the function 3dDeconvolve. The BOLD response to individual events from each stimulus condition was modeled using the sum of "TENT" basis functions – a piecewise linear spline function that estimates an impulse response function. The sum of 13 tent functions was used to cover the duration of 18 s after the stimulus onset ("TENT(0, 18, 13)"). All models were fit separately to each voxel. For each voxel within the pre-defined ROIs 13 amplitudes for each stimulus condition were estimated, which were the time course of the estimate hemodynamic response function

(HRF) to each condition at the 13 time points (from 0 to 18 s with the time step of 1.5 s). The mean of the first and last two time points was subtracted from each HRF to ensure that it started from and returned to approximately the same baseline level. Estimates from individual voxels were averaged within each of the 6 ROIs, and BOLD response amplitudes were estimated using the difference between the peak response (reached around 4.5 – 6 s after stimulus onset) and the baseline response. Our analysis focused on response differences between contour aligned and unaligned conditions, instead of direct comparison between aligned or unaligned contours in clutter versus not, to avoid confound of blood stealing from the background stimulus. Response differences between conditions were assessed using a bootstrapping procedure – resampling 12 subjects data with replacement within conditions and calculating differences across 10000 iterations, and a one-sided permutation test was performed to acquire *p*-values (Efron & Tibshirani, 1993). Analysis of variance (ANOVA) was also conducted to compare BOLD response amplitudes among conditions in each predefined ROI.

Connectivity analysis

Besides the estimated response to each stimulus condition, we also wanted to know whether interregional connections among these ROIs depend on the experimental conditions. Two functional connectivity analysis methods were used to answer this question. The first method, psychophysiological interactions (PPI), uses interaction terms created from the dot product of the seed region time series and vectors representing different experimental conditions to explain variance in time series from other cortical regions with a GLM analysis. If the estimated beta weights for the interaction regressors

are different among experimental conditions, the coordination between the test ROI and the seed ROI depends on the conditions. In the second method, beta series correlations, a beta weight for each experimental trial is estimated using GLM, these beta weights are sorted according to the experimental condition during the trial, and then correlations are calculated between the beta weights for each condition (Rissman, Gazzaley, & D'Esposito, 2004). If these correlations vary depending on the conditions, connectivity between the two regions is condition-dependent.

Psychophysiological interactions. First, the data modeled by the GLM (driving effect of stimuli) was subtracted from the pre-processed event-related dataset to generate a residual dataset for further analysis of the intrinsic interactions between cortical areas. Figure 3.3 shows an example of the PPI terms (regressors in a GLM). In order to build the psychophysiological interaction terms, both psychological condition codes (indicating the representation time for a certain stimulus condition) and physiological responses from the seed region (selected based on the BOLD response modulation across conditions) were required. According to McLaren, Ries, Xu, and Johnson (2012), with more than two conditions separately building one interaction term for each stimulus condition could be more robust to noise and have better model fits than directly using PPI terms with condition contrast (i.e., "1" for one condition while "-1" for a different condition). Therefore, we created one text file matching the length of functional time series for each condition as the condition code (D_c), in which "1" indicated stimulus presentation of that particular condition and "0" for the rest time points (Fig. 3B right panel). The physiological response from the seed region was estimated by deconvolving the time

series from the seed region with its HRF estimated from the previous analysis (Fig. 3A). Gitelman, Penny, Ashburner, and Friston (2003) and Kim and Horwitz (2008) demonstrated the importance of modeling the underlying neural activity: interaction should be expressed at a neuronal level rather than at the level of hemodynamic responses – neuronal activity being filtered with an HRF. Interaction was calculated as the product of condition codes and estimated physiological responses (Figure 3.3B). In order to compare the interaction with BOLD measurements, we then convolved the interaction at a neuronal level with HRFs to obtain a regressor at the level of hemodynamic responses (Figure 3.3C).

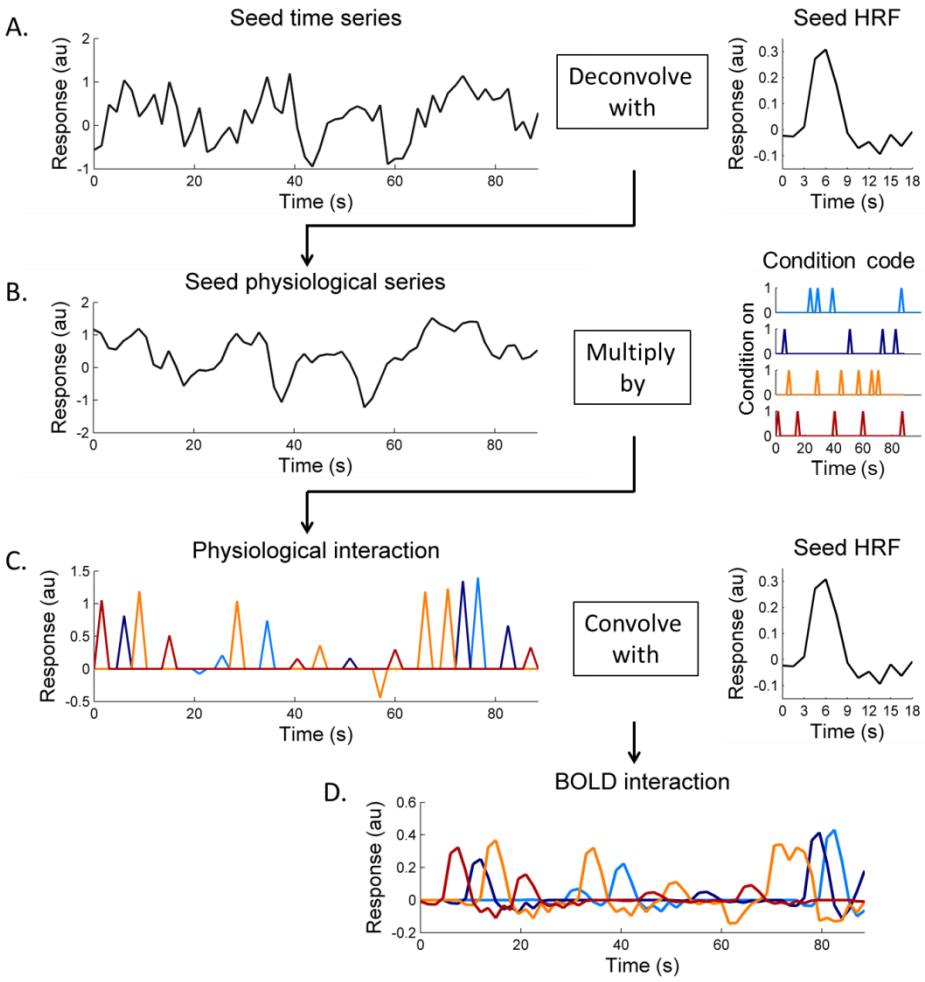

Figure 3.3 An example of the psychophysiological interactions term. A. The seed time series were first deconvolved based on the estimated HRF to obtain physiological responses. B. To combine both the physiological and psychophysical effects we multiplied the seed physiological series by the condition code from each condition separately. C. The interaction at the physiological level was convolved with the estimated HRF, so we could compare this BOLD level interaction (as shown in D) with residual time series from other ROIs.

In summary, residual time series from the *i*th ROI, y_i, could be modeled as (Friston et al., 1997):

$$y_i = \sum_{c=1}^{4}(\beta_{c,i} \cdot H(D_c \cdot n_{seed})) + \beta_{seed,i} \cdot y_{seed} + \varepsilon_i, \ (*)$$

where *c* is the stimulus condition index, here, we have four conditions, $c = 1, \ldots, 4$; $\beta_{c,i}$ is the correlation coefficient for the PPI regressor of condition *c* at ROI *i*; $\beta_{seed,i}$ is the correlation coefficient for the seed time series regressor at ROI *i*; $H(\cdot)$ indicates the convolution operation with the estimated TENT HRF for a certain condition (*c*) in the corresponding region (here, the seed region); D_c is the stimulus presentation code for condition *c*; y_{seed} is the time series from the seed region and n_{seed} is the physiological response estimated by deconvolving y_{seed} with the estimated HRF, that is, since $y_{seed} = H(n_{seed}) + noise$, we could solve for n_{seed} given the kernel function and y_{seed} using deconvolution; finally, ε_i is an error term at region *i*. The beta estimated for the PPI regressor represents the amount of signals could be explained by both the response in the seed ROI and the stimulus condition. If the beta estimates at a certain ROI from two conditions are different, the seed may differently influence this ROI between these two conditions. Next, 3dDeconvolve and 3dREMLfit (which estimates and removes noise temporal correlations) were used to estimate coefficients based on the model (*). Estimates of $\beta_{c \cdot i}$ from individual voxels were averaged within ROIs, and would be used for assessing statistical significance.

Beta series correlations. With the beta series correlation method (Rissman et al., 2004), we first estimated a beta weight for each experimental trial, that is, modeling the time series from the *i*th ROI, y_i, as

$$y_i = \sum_{c=1}^{4}\sum_{j=1}^{c_j}(\beta_{i,c,j} \cdot H_i(D_c)) + \varepsilon_i, \ (**)$$

where *c* is the index for the stimulus conditions; *i* is the index for the ROIs; $H_i(D_c)$ indicates the estimated hemodynamic response function at the region *i* for the condition *c*; $\beta_{i,c,j}$ is the beta weight for the *i*th ROI during the *j*th trial of the condition *c*; c_j is the total number of trials for the condition *c*; ε_i is an error term at region *i*. Next, the estimated beta values $\beta_{i,c,j}\ (j = 1, \dots, c_j)$ from the *i*th region were regressed against the $\beta_{i',c,j}$ from the region i'(the seed) according to the conditions, and the estimated linear coefficient was used to indicate the connectivity between regions i and i' under a certain experimental condition.

3.3 Results

The early visual areas V1, V2, and V3 were manually defined according to the polar angle (Figure 3.4A) and eccentricity phase maps acquired in separate scanning sessions. Functional localizers (three runs, pre-processed, averaged and subjected to Fourier analysis) were used to define ROIs corresponding to the cortical representations of stimulus target or background regions (Figure 3.4B and C). On the flat patch, a band of activation associated with the target representation (blue, Figure 3.4C) and two bands associated with the background representations (orange, Figure 3.4C) were consistent with the eccentricity features in the early visual areas. Therefore, two sets of ROIs were defined in each visual area corresponding to target (tgROIs) and background regions (bkgdROIs).

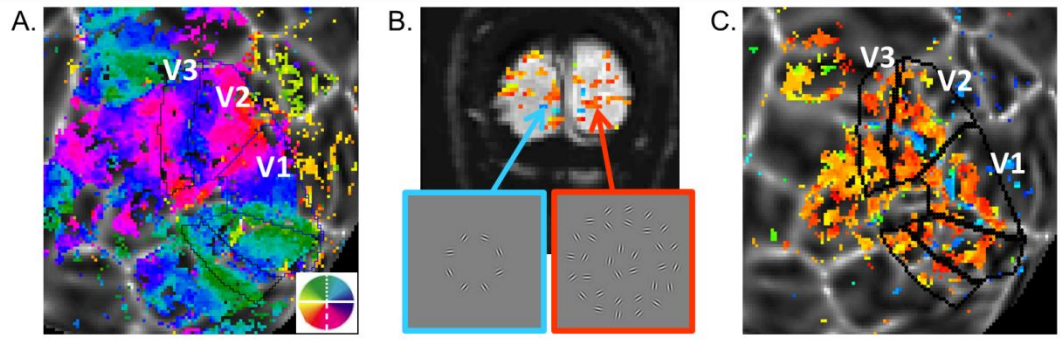

Figure 3.4 Visual area mapping (A) and functional localizer results (B and C). A. Angular visual field preference of one observers' left hemisphere obtained from rotating wedge stimulus (overlay on a flattened patch of the cortical surface centered on the occipital pole). The early visual areas are labeled. B. On one single coronal EPI image, voxels significantly correlated with the block-alternation are color coded based on relative phases – the bluish voxels are in phase with the target presentation, while the orange voxels are in phase with the background stimulus. C. Data in B was transformed to the flat patch, where a blue target-associated band and two orange background-associated bands could be seen among the early visual areas.

3.3.1 Stimulus-related activity

We first looked at the BOLD response magnitude for each experimental condition. Estimated HRFs in one representative subject within the tgV2 ROI are shown in Figure 3.5. The differences of BOLD responses to the aligned from the unaligned contours for each ROI depend on the background context as shown in Figure 3.6. When the background was present, the tgV2 ROI showed a significant preference for the aligned contours ($\mathrm{albg} - \mathrm{uabg} = 0.057$, with one-sided permutation test $p < 0.001$). In tgV1 the responses to aligned contours were weaker than the responses to the unaligned contours ($\mathrm{alnb} - \mathrm{uanb} = -0.049$, with one-sided permutation test $p = 0.036$) when

there was no background. Using a two-way ANOVA model within each visual area (Alignment and Background as fixed effects, and subjects as a random effect), a significant interaction between Alignment and Background was observed in tgV2 ROI ($F[1,47] = 5.63, p = 0.037$) and a similar trend was seen in tgV1 ROI ($F[1,47] = 4.72, p = 0.053$). When we tested for this interaction using a permutation test, tgV2 showed $p = 0.0089$, and tgV1 showed $p = 0.011$. No significant effect was found in background ROIs. Thus, the regions in V1 and V2 corresponding retinotopically to the target ring responded to the coherent contour differently according to the context.

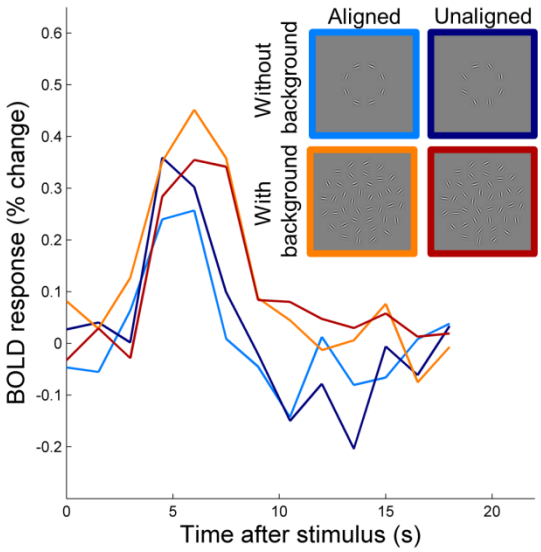

Figure 3.5 Estimated HRFs from four experimental conditions for one subject's tgV2 ROI. The HRFs were estimated for each voxel, and then averaged within each ROI.

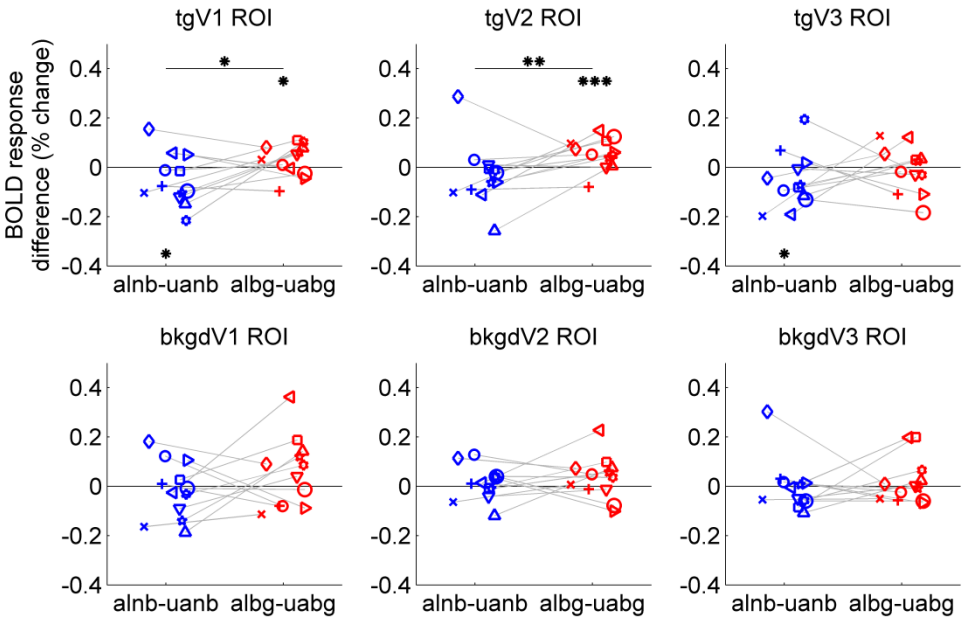

Figure 3.6 Stimulus-related BOLD response differences among conditions. Each panel shows data from one ROI for individual subjects (shown in different icons). Blue icons are differences of estimated HRF amplitudes between the aligned and unaligned contours when there was no background, and red icons are BOLD differences when there was background clutter. Gray lines connect data points from the same subject. Asterisks indicate statistical significance based on one-sided permutation test at *$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$.

3.3.2 Coordination among regions influenced by physiological and psychological states

We also explored coordination among defined ROIs using two connectivity analysis methods. With the PPI analysis, we used the tgV2 ROI, which showed strong modulations among conditions, as the seed region. The PPI connectivity results are shown in Figure 3.7. When the background was present, beta estimates for the aligned

condition were larger than the unaligned ($\mathrm{albg} - \mathrm{uabg} = 0.18$, with permutation test $p = 0.0085$) in tgV1 ROI. Using a two-way ANOVA model within each ROI (Alignment and Background as fixed effects, and subjects as a random effect), a significant interaction between Alignment and Background was observed in tgV1 ROI ($\mathrm{F}[1,47] = 5.38, p = 0.040$). The correlation differences were also retinotopically specific to the target ROIs: no effect was observed in the background ROIs.

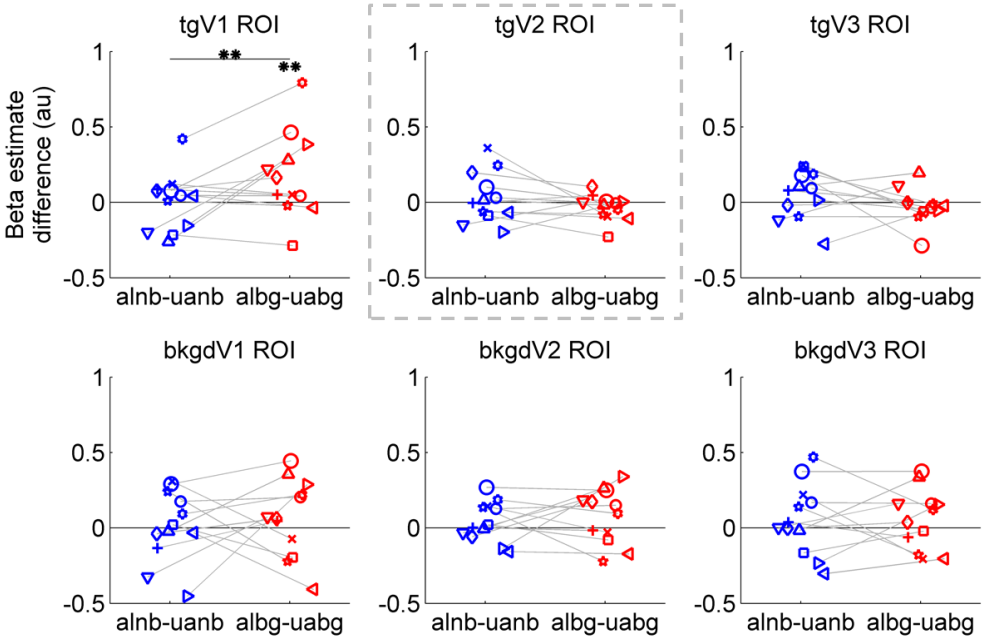

Figure 3.7 Connectivity results using PPI. Differences of Beta estimates among conditions are shown when tgV2 was the seed (dashed frame). Blue icons are differences of estimated beta weights of PPI terms between the aligned and unaligned contours when there was no background, and red icons are differences when there was background clutter. Gray lines connect data points from the same subject. Asterisks show significant levels based on the permutation test.

With a second inter-area connectivity analysis, the beta series correlations, a beta value was first estimated for each experimental trial. Next, these beta values were sorted according to conditions and correlated among ROIs for each condition (Figure 3.8A). Figure 3.8B shows differences of beta series correlations estimated in tgV1 against tgV2 ROI. With the background, the linear coefficients of beta values from the tgV1 against beta values from the tgV2 were larger when the Gabors were aligned ($\mathrm{albg} - \mathrm{uabg} = 0.12$, with one-sided permutation test $p < 0.001$); a weak trend of interaction between the Alignment and Background was also observed (with permutation-based $p = 0.045$; ANOVA, $F[1,47] = 2.54, p = 0.14$). The ANOVA *p*-values of the Alignment and Background interactions are shown in Figure 3.8C, which shows similar sensitivity to experimental conditions of tgV1-tgV2 connectivity as using the PPI analysis. A significant interaction was also observed in tgV3 when tgV2 was the seed ($p = 0.0058$), and in bkgdV3 when bkgdV2 was the seed ($p = 0.022$).

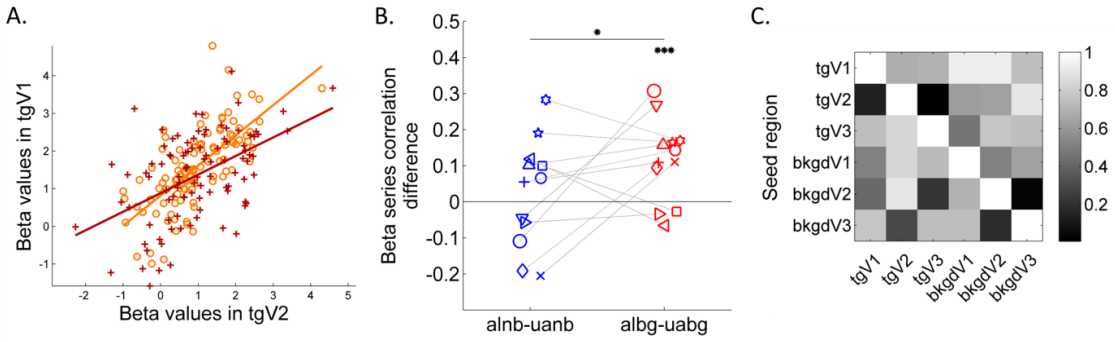

Figure 3.8 Connectivity results using beta series correlations.A. Scatter plot of the beta values in tgV1 against those in tgV2 for a single subject under the albg (in orange) and uabg (in dark red) conditions. B. Differences of beta series correlations estimated in tgV1 against tgV2 ROI for individual subjects (shown in different icons). Asterisks show

significant levels based on the permutation test. With background, the aligned condition tended to show larger linear coefficients between the tgV1 and tgV2. C. The *p*-values from the interactions between the Alignment and Background factors (ANOVA). A small *p*-value is shown in dark color. Row indicates the seed region. When tgV2 is the seed, this interaction in tgV1 and tgV3 is significant; when bkgdV2 is the seed, a significant interaction is observed in bkgdV3.

3.4 Discussion

3.4.1 Enhancing and suppressive responses in early visual areas

We investigated how context and contour coherence affect the magnitude and inter-area correlations of the fMRI BOLD signal in human early visual areas. The results from our first analysis showed that with clutter in the background, target ROIs in early visual areas V1 and V2 had preference to the aligned contours (Altmann et al., 2003; Kourtzi et al., 2003); while with isolated structure, preference for the unaligned contours was observed (Murray et al., 2002). This interaction between the contour alignment and the background context was significant in the regions retinotopically corresponding to the target stimulation in areas V1 and V2. Area V2 showed the most significant effect, which agrees with its role in extracting features from complex visual scenes (Boynton & Hegdé, 2004; Huang, Hess, & Dakin, 2006; Minami Ito & Komatsu, 2004; Merigan, Nealey, & Maunsell, 1993; Anna W. Roe, 2003; von der Heydt, Peterhans, & Baumgartner, 1984). In contrast to previous functional imaging studies, we used a separate functional localizer to define the cortical region retinotopically associated with the target contour; and we demonstrated that the effect was specific to the target related region, rather than a diffuse

effect among early visual areas, since no significant difference was observed in the background ROIs.

Our results are consistent with the electrophysiological results reported in W. Li et al. (2006). They showed a close correlation between the responses of monkey V1 neurons and the perceptual saliency of contours which was modulated by number of collinear elements or relative spacing between them; however, the correlation could be either positive or negative, depending on the context beyond the collinear elements. Specifically, they found that without the background clutter, neurons in V1 showed facilitation to three collinear lines compared with a single line in their receptive fields; but by adding more aligned line segments, responses in V1 neurons turned to be inhibited. However, with the background clutter, neuronal responses increased monotonically with increasing aligned line segments.

We found both facilitatory and suppressive results at the target ROIs in a contour detection task. These results suggest that both local feature-guided and global form-guided strategies could be used in contour integration, and they interact according to the context. The facilitatory results could be predicted based on flank facilitation on local segments along the contour. Both psychophysics and electrophysiology has shown that when surround segments were positioned within a certain range outside the neuron's receptive field and placed collinearly with the central stimulus, the neuronal responses to the center would be facilitated (C.-C. Chen & Tyler, 2008; Kapadia et al., 1995; Kapadia et al., 2000; Polat & Sagi, 1993, 1994). One interpretation is that the intrinsic horizontal connections in V1 can link neurons with nonoverlapping receptive fields but with similar

orientation preference to integrate information over a relatively large visual field (Angelucci et al., 2002; Bosking, Zhang, Schofield, & Fitzpatrick, 1997; Gilbert, 1992; Gilbert & Wiesel, 1989; Z. Li, 1998; Malach, Amir, Harel, & Grinvald, 1993; McGuire, Gilbert, Rivlin, & Wiesel, 1991; Rockland & Lund, 1983; Stettler, Das, Bennett, & Gilbert, 2002; Ts'o, Gilbert, & Wiesel, 1986). Therefore, the facilitated single neuron responses can be associated to form a coherent contour (Field et al., 1993; Hess & Field, 1999; W. Li & Gilbert, 2002), and cause response increases along the path. Furthermore, feedback signals from higher cortical areas, such as area V4, can also enhance the global contour signals in early visual areas (M. Chen et al., 2014).

In contrast, the suppressive results rely more on an understanding of the global scene. The visual system is wired to efficiently represent structure following natural scene statistics (Attneave, 1954; Barlow, 1961; Simoncelli & Olshausen, 2001). One way to ensure the efficiency is to generate high-level "summary" templates for probable forms of the natural inputs, such as cocircularity (Sigman et al., 2001), and only signals representing deviations from the predicted templates are carried forward to be resolved by further processing, a theory referred to as "predictive coding" (Friston, 2005; MacKay, 1956; Mumford, 1992; Murray et al., 2004; Rao & Ballard, 1999). Increases or decreases in deviations would result in a change in localized neural activity (but see de-Wit et al. (2012)). Alternatively, the predictions from a higher level could disambiguate the lower-level representation by attenuating responses to unmatched incoming features (Murray et al., 2004; Yuille & Kersten, 2006). For example, the signal to background clutter could be suppressed, and this may decrease overall averaged responses to the

condition with coherent structure, perhaps for the purpose of metabolic efficiency (Barlow, 1959, 1961). Based on this model, when a coherent target appears, cortical responses to the target would be enhanced whereas responses to surrounding noise would be suppressed, as shown by Gilad et al. (2013) using voltage-sensitive dye imaging in V1 of monkeys. Based on our results, the suppression was only observed at the regions specifically respond to the target ring, and this is consistent with the predictive coding idea. No BOLD response suppression in background ROIs was observed, which could be due to the complexity of BOLD signals, especially from neural inhibition (Buzsáki, Kaila, & Raichle, 2007; Lee et al., 2010; Schumacher & Olman, 2010).

3.4.2 Coordination between areas V1 and V2

Furthermore, using the connectivity analysis – the psychophysiological interactions and beta series correlations – we found that the coordination between target V1 and target V2 ROIs was also highly dependent on the stimulus conditions. While the contours were presented together with the background clutter, a larger connectivity was observed between tgV1 and tgV2 ROIs for aligned contours than for unaligned ones, which could be predicted from the theoretical framework in areas V1 and V2 by Roelfsema, Lamme, and Spekreijse (2000); however, no connectivity difference was observed when the contours were presented without the background.

Close connections between areas V1 and V2 are well established from anatomical and physiological studies. A large percentage of the cortical inputs in visual area V2 is from area V1 (Felleman & Van Essen, 1991; Sincich, Adams, & Horton, 2003), and they form multiple parallel pathways each carrying specific local representations such as

color, form and motion from V1 neurons (Federer et al., 2009; Livingstone & Hubel, 1988; Sincich & Horton, 2002, 2005). Many V1 neurons that project directly to V2 are tuned for orientation, which plays an important role in solving for global contours (El-Shamayleh, Kumbhani, Dhruv, & Movshon, 2013). It has been demonstrated that these feedforward connections are critical for responses in V2 neurons, whose activity is abolished when V1 is inactivated by cooling (Girard & Bullier, 1989; Schiller & Malpeli, 1977). Area V1 also receives numerous feedback projections from V2 (J. C. Anderson & Martin, 2009; Angelucci et al., 2002; Barone, Batardiere, Knoblauch, & Kennedy, 2000; Girard, Hupé, & Bullier, 2001; Rockland & Virga, 1989; Stettler et al., 2002), but their function could be highly dependent on the visual stimulus (J. C. Anderson & Martin, 2009). For example, though with one isolated stimulus no change in orientation selectivity was observed in area V1 when area V2 was cooled (Sandell & Schiller, 1982) or with simple center-surround stimulus only a few V1 neurons were affected by V2 inactivation (Bullier, Hupé, James, & Girard, 1996; Hupé, James, Girard, & Bullier, 2001), it is possible that feedback plays a stronger role in more complex scenes. For example, given illusory contours induced by abutting gratings, there is evidence that area V2 modulates the orientation representation map in area V1 to provide a signature for a "higher order" contour (Ramsden, Hung, & Roe, 2001; Anna W. Roe, 2003). The third category of connections that would relate responses in area V1 to V2 is the common inputs to these two areas from the same cortical and subcortical structure (Kennedy & Bullier, 1985). One example is that responses in areas V1 and V2 are strongly influenced by feedback function from area MT (Dakin, 2009; Hupé et al., 1998; Sillito, Cudeiro, &

Jones, 2006). Though here we are not able to distinguish effects among the above three connections, the current results showed that the coordination between area V1 and V2 did rely on the contour alignment and the context.

3.4.3 Importance of flexibility

Overall, the finding that the responses of cortical regions retinotopically corresponding to coherent contours could be facilitated or inhibited (or the inter-area correlation could increase or decrease) depending on the scene context implies more than one strategy in contour processing. If a single strategy is being used, say, increasing responses to coherent contours, we should expect the response increase to the aligned contours for both with and without background conditions. However, this was not what we observed. In our case, with background, the target tends to be buried in noise, and enhancing the target signal could be necessary for the system to accurately separate the target from the background (Lamme, 1995; Self, van Kerkoerle, Supèr, & Roelfsema, 2013). However, without background, once the system figures out representations of a coherent circular contour, no forwarding "error" term (the discrepancy between the lower level inputs and the internal template) is necessary, which would lead to a response decrease to the coherent target. Here, the system could be seeking to encode information economically to reduce representations of redundant information (Attneave, 1954; Barlow, 1961). In terms of the coordination among areas, it is possible that with the background or distractors, the inter-area correlation between areas V1 and V2 is dominated by concurrent signal transferring between area V1 to V2 to isolate the target contour from background noise. The similar inter-area correlation increases with

increasing levels of elements alignment or recognizable features have been observed in previous studies (Cardin et al., 2010; M. Chen et al., 2014; J. Freeman, Donner, & Heeger, 2011). However, when there is no background clutter, this constant signal feeding may be interrupted by higher level functions for the purpose of efficiency. Overall, balancing the above two strategies according to a greater stimulus context could improve both accuracy and efficiency of the cortical function.

Computationally, information could be grouped based on either global or local features. The strategy for information integration relying on the global features should be different from the strategy for information integration relying on local linkage of similar features. The former is more top-down: as in Figure 3.9A, a coherent circular shape could be perceived even when individual elements are inconsistent. In contrast, the latter strongly depends on bottom-up relays, and it relies on similarities among nearby elements (Figure 3.9C). Among clutter, when the nearby elements forming the circle do not share features, the integration process would fail (Figure 9B; see also Keeble and Hess (1999), Levi and Klein (2000)). Therefore, constant bottom-up inputs from early visual regions linking similar low-level features are necessary in a contour integration task among clutter. In fact, the prediction of human performance for detecting naturalistic contours among background distractors is fairly accurate by applying local grouping functions (Geisler et al., 2001). Additionally, due to potential ambiguities from merely bottom-up processes, a higher level template could be useful (Elder, Krupnik, & Johnston, 2003; Epshtein, Lifshitz, & Ullman, 2008). In all, applying both strategies can be crucial in visual processing (Friston, 2005; Kersten, Mamassian, & Yuille, 2004; Kersten & Yuille,

2014), and the system would be able to dynamically adjust efficacy given the signal-to-noise ratio in the scene (Zhaoping, 2014).

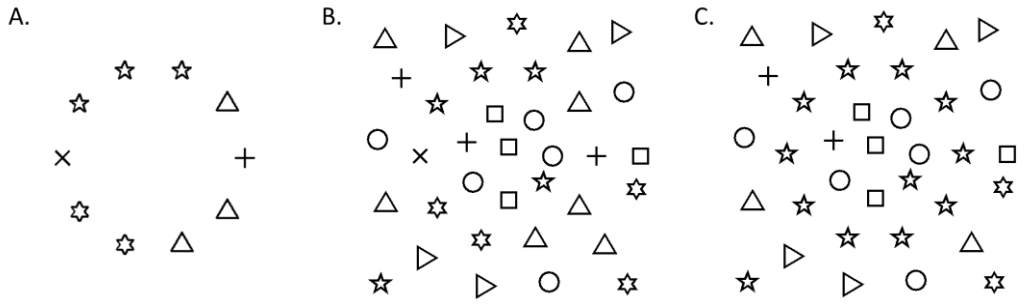

Figure 3.9 Local linkage cues are required with background clutter. A. A circular shape could be grouped based on global features. B. The same circle is not easy to be identified when surrounded by clutter. Local linkage cues are required. For example, in C, the nearby pentagrams serve as local cues for the circular shape to be grouped.

3.5 Conclusion

In summary, we have shown that the cortical responses of human early visual areas to coherent contours are affected by a larger context around them. A coherent target enhances neuronal responses to indicate certainty, but this may not be efficient especially when the target could be easily abstracted or explained. On the other hand, a system that only relies on the predictive coding approach and entirely discards basic representation signals from lower level areas is inflexible and may encounter problems later on, e.g., when further operations on other detailed features are requested. Therefore, the visual system should be able to apply both strategies and to weight them according to context, such as scene complexity or task difficulty. With the current task, detecting circular contours among simple or complex scenes, we have found that early visual areas V1 and

V2 may play an important role in manipulating contour integration strategies under various conditions.

Acknowledgements

The authors would like to thank D. Do, J. F. Schumacher, A. Grant, and M-P. Schallmo for scanning assistance. This work was supported by funding/grant sources: NIH R21NS075525, P30NS076408, P41 EB015894, and ONR N000141210883.

4. Segmentation decreases the magnitude of the tilt illusion[1]

Qiu, C., Kersten, D., & Olman, C. A. (2013). Segmentation decreases the magnitude of the tilt illusion. *Journal of Vision*, *13*(13), 1–17. http://doi.org/10.1167/13.13.19

Summary

In the tilt illusion, the perceived orientation of a target grating depends strongly on the orientation of a surround. When the orientations of the center and surround gratings differ by a small angle, the center grating appears to tilt away from the surround orientation (repulsion), whereas for a large difference in angle, the center appears to tilt toward the surround orientation (attraction). In order to understand how segmentation/perceptual grouping of the center and surround affect the magnitude of the tilt illusion, we conducted three psychophysical experiments in which we measured observers' perception of center orientation as a function of center-surround relative contrast, relative disparity depth, and geometric features such as occlusion and collinearity. All of these manipulations affected the strength of perceived orientation bias in the center. Our results suggest that if stronger segmentation/perceptual grouping are induced between the center and surround, the tilt repulsion bias decreases/increases. A grouping-dependent tilt illusion plays an important role in visual search and detection by enhancing the sensitivity of our visual system to feature discrepancies, especially in relatively homogenous environments.

[1] The Association for Research in Vision and Ophthalmology (©ARVO) as the copyright holder

Keywords: tilt illusion, segmentation, perceptual grouping, human psychophysics

4.1 Introduction

Many visual illusions are the result of contextual modulation: influenced by contextual information, we often perceive things differently from their physical reality. In the case of orientation perception, it has been demonstrated that the orientation of the surround affects the perceived orientation of the center (Blakemore, Carpenter, et al., 1970; Gibson & Radner, 1937; Goddard, Clifford, & Solomon, 2008; Schwartz, Sejnowski, & Dayan, 2009). A central grating is perceived as tilted away from the orientation of a surround grating when the two orientations are similar; this is called the direct (repulsion) form of the tilt illusion. When the center and surround orientations differ considerably, the perceived orientation of the central grating is attracted toward the surround orientation, which is known as the indirect (attraction) tilt illusion. The relative orientation between the center and surround determines whether we perceive the repulsion or attraction effect.

The neural basis for the tilt illusion can be modeled as changes in the tuning curves of individual orientation-selective units in the presence of the surround (Blakemore, Carpenter, & Georgeson, 1971; Blakemore & Tobin, 1972; Clifford, Wenderoth, & Spehar, 2000; Gilbert & Wiesel, 1990; Schwartz, Hsu, & Dayan, 2007), and with the perceived orientation of the center being determined by the vector average (Georgopoulos, Schwartz, & Kettner, 1986) of the units' responses. The effect can also be modeled by lateral interactions at the population level (Bednar & Miikkulainen, 2000; Solomon, Felisberti, & Morgan, 2004). Electrophysiological results have demonstrated

that modulations of neural response by surrounding context include magnitude variation (Cavanaugh et al., 2002b; Levitt & Lund, 1997; W. Li et al., 2000; Muller, Metha, Krauskopf, & Lennie, 2002; Sengpiel, Sen, & Blakemore, 1997; van der Smagt, Wehrhahn, & Albright, 2005), broadening or sharpening of tuning widths (Gilbert & Wiesel, 1990), and repulsive or attractive shifts in preferred orientation (Felsen, Touryan, & Dan, 2005; Gilbert & Wiesel, 1990). The tuning curve changes may serve to optimize sensory coding (Clifford et al., 2000; Schwartz et al., 2007; Simoncelli, 2003). Using principles of efficient coding of the input signals, the extra constraint provided by the context allows the central detectors to remove statistical dependencies, which acts as a transform that reduces redundancies among inputs (Attneave, 1954; Barlow, 1961; Z. Li & Atick, 1994; Olshausen & Field, 1996). A simple efficient coding transform is divisive gain control normalization (Albrecht & Geisler, 1991; Carandini & Heeger, 1994; Carandini & Heeger, 2012; Heeger, 1992; Lyu, 2010, 2011), which nicely explains nonlinear response properties of neurons in primary visual cortex (Carandini, Heeger, & Movshon, 1997; Schwartz & Simoncelli, 2001; Simoncelli & Schwartz, 1999).

Further studies (Coen-Cagli, Dayan, & Schwartz, 2012; Schwartz, Sejnowski, & Dayan, 2006; Schwartz et al., 2009) have shown that this divisive normalization process may only apply when the center and context are perceptually assigned to the same object or segment in natural scenes. Based on statistical measurements in natural images, Schwartz et al. (2009) reported that across segmentation boundaries, the orientation dependence between the central and surround patches was greatly reduced. Therefore, they proposed to combine a segmentation factor with a divisive gain control model to

account for natural image statistical dependence more accurately. This model provides a unified explanation for both repulsion and attraction in the tilt illusion. In their model, the segmentation factor is controlled by center-surround relative orientations. The closer the relative orientations, the more likely they share the same gain pool. In this study, we explored whether segmentation/perceptual grouping cues other than relative orientation could be used by the visual system in a similar way to manipulate the tilt effects. Specifically, we tested local image features of relative contrast and disparity depth (Exp. 1), and geometric features, such as occlusion (Exp. 2) and collinearity (Exp. 3), in influencing the perception of central orientations.

For relative contrast, the greatest tilt repulsion occurs when the center and surround gratings have the same contrast (Durant & Clifford, 2006; Tolhurst & Thompson, 1975), suggesting that contrast differences might provide segmentation cues to reduce the magnitude of the tilt bias. The effect of contrast cues on the tilt attraction has not been studied thoroughly. Previous findings showed that manipulations of spatial separation or spatial frequency have no significant effect on the attraction (Wenderoth & Johnstone, 1988) even though they change the tilt repulsion (Georgeson, 1973; Tolhurst & Thompson, 1975). We asked whether this pattern of results extends to the contrast cues, and whether the contrast cues affect the tilt repulsion and attraction differently.

Similarly, effects of depth disparity on the tilt illusion are unclear. Using line segments, Sakai and Hirai (2002) and Westhhmer (1990) observed that apparent tilt did not depend on the stereo disparity cues between the target and contextual bars. However, Durant and Clifford (2006) obtained reduction of the tilt repulsion with stereo disparity

cues between the center and surround gratings. Since relative depth cues, just like relative contrast cues, would influence perceptual segmentation between the center and surround, we expected that both could manipulate tilt biases (Exp. 1), in a manner that could be predicted by the Schwartz model (Schwartz et al., 2009).

In addition to the relative contrast and disparity depth between the center and surround, geometric features could also be an important factor for segmentation/perceptual grouping. For example, in three-dimensional space, an occluding ring in front of a border between target and context would make the border ambiguous, and actually encourage the grouping of the center and surround. This would leave the filtering input unchanged in the Schwartz model, and more directly reveal the effect of segmentation/perceptual grouping on the tilt repulsion (Exp. 2). Spatial layout is another cue that influences perceptual organization. Based on the natural image statistics, the end position along the central elements follows the most frequent direction of edge co-occurrence (Geisler et al., 2001; Sigman et al., 2001). This result indicates that surround patches that are collinear with a central grating would provide stronger evidence for contour grouping than patches flanking the center. Hence, we predicted that this collinear layout would show different effects on the central orientation perception (Exp. 3).

In the three experiments described below we systematically measured the tilt illusion affected by different segmentation cues between the center and surround, and sought to understand how the visual system responds to central orientation given various combinations of segmentation/perceptual grouping cues. In the first experiment, we

examined whether the tilt illusion could be influenced by contrast and depth differences between the center and surround. To account for our results, we expanded the segmentation model by Schwartz et al. (2009) to include the contrast and depth cues. We showed that the model could account for the decrease of the tilt effects and orientation-tuning shift of the tilt biases as a function of relative orientation. In the second experiment, we used an occluding ring to affect perceptual grouping while maintaining the filtering activation in the Schwartz model. As predicted, increase of perceptual grouping cues led to stronger tilt repulsion. In Exp. 3, we measured the tilt repulsion with different surround spatial layouts, which showed that the maximal repulsion bias occurred when gratings were along the end locations of the central stimulus.

4.2 Experiment 1: Relative contrast and depth

4.2.1 Methods

General. Six observers (mean age: 29 years, 3 males) with normal or corrected to normal visual acuity were tested. All were trained for a short time (2-5 min) in order to get acquainted with the task and to obtain ranges of individual stimuli variation. Four observers participated in all experimental conditions, and two observers completed six out of eight conditions.

Visual stimuli, sinusoidal gratings of the same mean luminance as the background, were generated using Matlab (Mathworks, Inc., Natick, MA) in conjunction with the Psychophysics toolbox (Brainard, 1997; Pelli, 1997). They were displayed on a high-resolution monitor (1600×1200 pixels, 60 Hz refresh rate, NEC MultiSync LCD 2190 uxi) connected to a Mac mini. Observers were seated 60 cm away from the screen.

A stereoscope and split screen were used in all conditions to maintain consistency across conditions, and at the beginning of each session, the stereoscope was adjusted by aligning two short nonius lines.

Procedure. Observers were shown stimuli at the center of the visual field, and were required to make binary judgments about the orientation of the central grating as tilted clockwise or counterclockwise from vertical. Stimulus duration was 500 ms. The observer's keyboard response initiated the next trial. A fixation point was displayed at the center of the screen at all times. The central circular test grating was 1 degree of visual angle in diameter and the surrounding annular grating was 3 deg in diameter. Both central and surround grating had a spatial frequency of 2 cycles per degree.

In order to obtain a psychophysical measure of subjective vertical, the orientation of the central grating was varied around the vertical based on a random, double staircase-method (Cornsweet, 1962). Subjective vertical in each condition corresponded to 50% of the clockwise (right-tilt) responses as estimated from the psychometric function, which was fit using the psignifit toolbox version 2.5.6 for Matlab (see http://bootstrap-software.org/psignifit/) implementing the maximum-likelihood method described by Wichmann and Hill (2001). For each center and surround condition, the tilt bias was defined as the subjective vertical difference between perceived orientation of the center with and without surround. Thus, the tilt bias eliminated any individual biases in orientation perception.

Eight viewing conditions (Figure 4.1) were employed to investigate three factors on the tilt illusion: contrast of the center grating, relative contrast, and relative depth

between the center and surround. In the conditions with depth difference, the surround appeared farther from the observer than the center, which appeared at the same depth through all conditions. Data from 16 relative center-surround orientations (ranging from 0 deg to 90 deg) were collected in each condition in order to characterize the tilt repulsion and attraction effect as a function of relative orientations. Measuring tilt bias across the entire range of relative orientations allows us to monitor the true features of repulsion and attraction, which might be missed by only recording observations from one relative orientation between the center and surround (such as 20 deg or 70 deg).

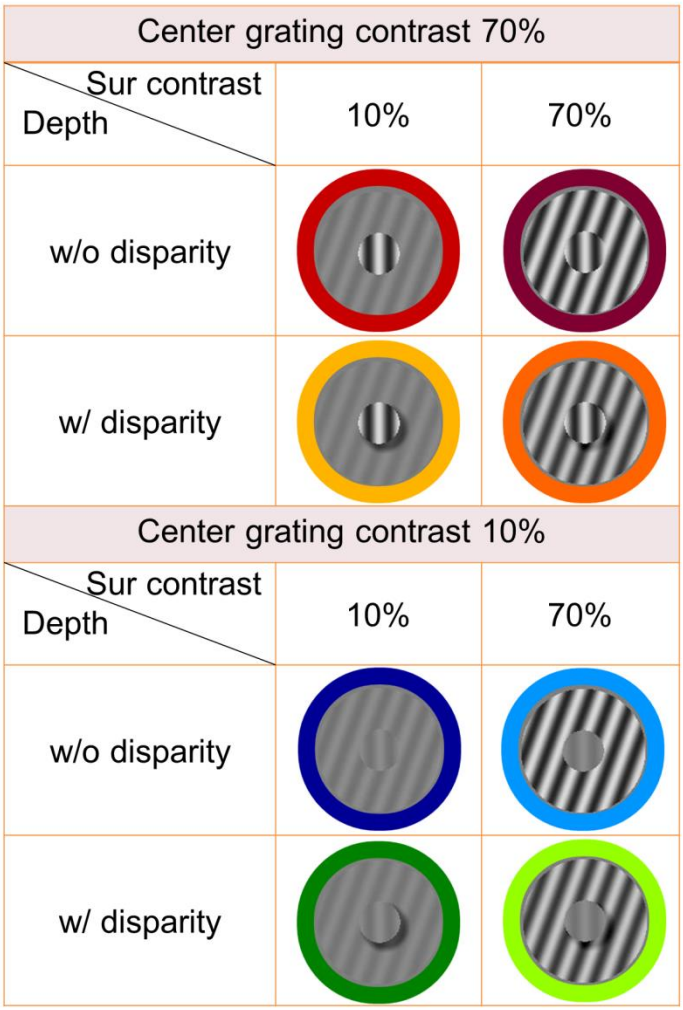

Figure 4.1 Example stimuli from eight conditions in Exp. 1. The darkest red and blue indicate the condition without any extra segmentation cues, while the yellow and light green show the conditions with both relative contrast and disparity cues. The disparity cue is illustrated using shadows. The conditions with reddish color code all have a high-contrast (70%) center, and bluish ones have a low-contrast (10%) center.

The subjective vertical under each condition for one of the 16 relative center-surround orientations was estimated in sessions of 80 trials. The sixteen estimates were

fit to a natural cubic smoothing spline with seven effective degrees of freedom. Five features were extracted from the fitted curve to quantify the repulsion and attraction effects (Figure 4.2): maximum repulsion, relative orientation at maximum repulsion, maximum attraction, relative orientation at maximum attraction and cross-over point (where the repulsion switches to the attraction effect). In Figure 4.3A, we report the mean of these features in six observers for eight conditions and the standard errors of these means. Analysis of variance (ANOVA) was used to compare among these group means (Figure 4.3B-D). Those sixteen estimates for each condition were also fit by a modified Schwartz model (Figure 4.4) to explore the relationship between the tilt effects and the segmentation features in different conditions.

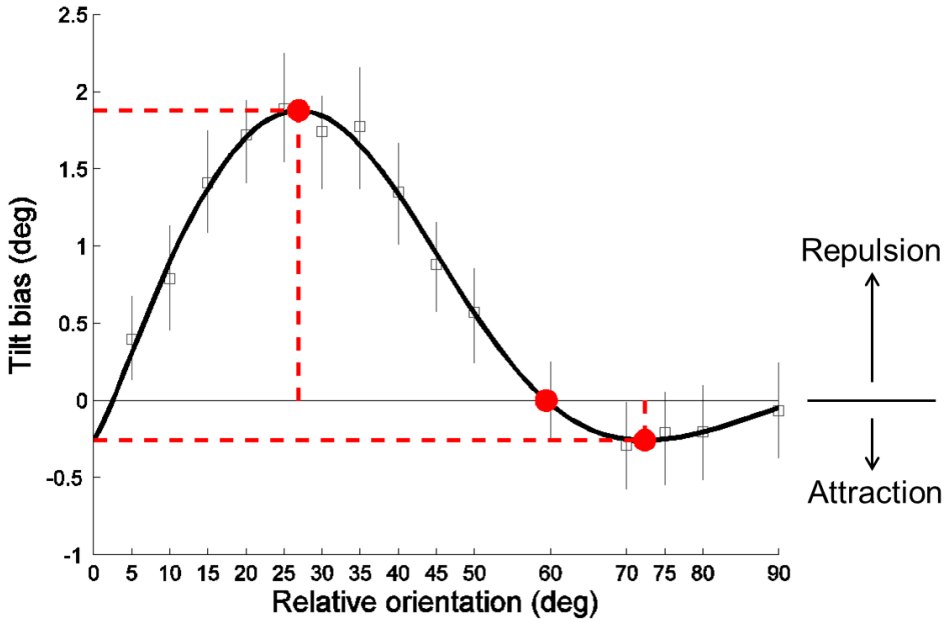

Figure 4.2 Example of a smoothing spline fit to data from one observer in one of the experimental conditions. The x-axis indicates the relative orientation between the center and surround. Sixteen relative orientations were sampled for each condition. The y-axis

indicates the tilt bias from the vertical. Positive biases indicate the repulsion effect, and negative biases indicate the attraction effect. Squares and error bars show tilt biases estimated based on individual psychometric functions. The solid black line shows the fitted curve. The red points represent features extracted from the fitted curve. In this example, the maximum repulsion is 1.88 deg when the center-surround relative orientation is 26.9 deg, the maximum attraction is 0.260 deg when the center-surround orientation is 72.4 deg, and at a relative orientation of 59.5 deg, the repulsion switches to the attraction effect (cross-over point).

4.2.2 Results

Figure 4.3A shows average repulsion and attraction peaks from eight experimental conditions. The points on the left show repulsion features, and the points on the right show attraction. Conditions with no contrast or depth segmentation cues tend to have stronger repulsion and attraction effects (dark red and blue points). When contrast of the center grating is high (reddish points), tilt biases (the repulsion and attraction) were reduced by either contrast cues, depth cues or both. The attraction effect almost vanishes in the high-contrast center condition with both depth and contrast cues (yellow points). Conditions with a low-contrast center grating (bluish points) show great variation in terms of the attraction effect. The conditions with the low-contrast center and high-contrast surround (light green and blue points) show stronger attraction, even when presented with contrast or depth segmentation cues.

A mixed-effects model was used in an ANOVA to clarify fixed effects of central contrast, center-surround relative contrast and depth, and their interactions, while considering subject as a random effect. This test was performed for the five extracted

features separately (see Appendix, Table A1). The factors of contrast ($F(1, 33) = 58.7, p < .001$) and depth ($F(1, 33) = 21.3, p < .001$) are significant in manipulating the maximum repulsion (Figure 4.3B): both segmentation cues reduce the repulsion effect, whereas perceptual grouping cues increase the effect. For maximum attraction (Figure 4.3C), there is a strong interaction between the central contrast and the relative contrast cue ($F(1, 33) = 22.1, p < .001$). The conditions with a low-contrast center but high-contrast surround (with relative contrast cue) show much stronger attraction. With regard to the cross-over points (Figure 4.3D), when the contrast of the center grating is high but surround contrast is low, the range of repulsion is much greater ($F(1, 33) = 23.2, p < .001$). This is consistent with a much weaker attraction effect in these conditions. In addition, the presence of depth cues significantly enlarges the range of repulsion ($F(1, 33) = 7.84, p = .008$).

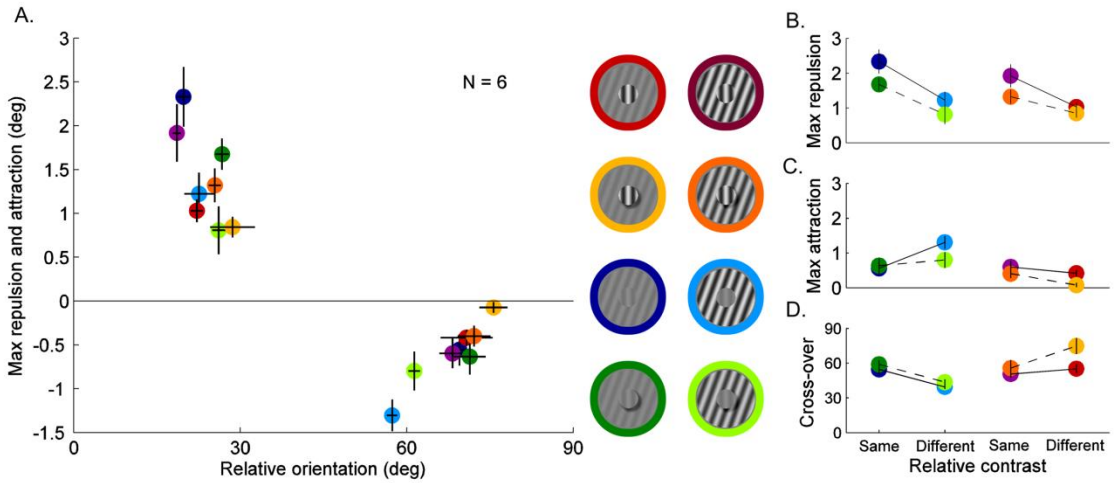

Figure 4.3 Results from Exp. 1. A. Points are peaks (maximum repulsions) and valleys (maximum attractions) of the smoothing spline curves from various conditions. Error bars are ±1 *SE* of maximum bias (vertical) or its corresponding

relative orientation (horizontal). (B-D) Results plot based on factors. The bluish points represent conditions with low central contrast, and the reddish points represent conditions with high central contrast. Solid lines connect conditions with no stereo disparity cue, while dash lines connect conditions with disparity cue. The x-axis shows center-surround relative contrast cues, either the same or different as shown. B. The maximum repulsions from eight conditions. C. The maximum attractions (the absolute values). D. The cross-over orientations extracted from the smoothing spline curves.

4.2.3 Model

The results described above suggest that contextual cues, such as relative contrast and depth, can affect the perception of tilt. In order to better understand the psychophysics of the perceived orientation changes in our results, we used a computational model to relate the role of context in perceived orientation to neural activity. Specifically, a computational model proposed by Schwartz et al. (2009) treats relative orientation as a cue to probabilistically co-assign the center and surround in the gain pool, and then a divisive gain control process combined with this co-assignment probability could well explain the tilt repulsion and attraction. Here, we considered whether this model could be expanded to include additional segmentation cues and predict our results.

The Schwartz model has two main components to describe center/surround interaction: divisive normalization and segmentation. Divisive normalization can serve to reduce redundant information, for example, the orientation dependence between the

center and surround in natural scenes (Schwartz & Simoncelli, 2001; Simoncelli & Schwartz, 1999; Valerio & Navarro, 2003). However, increased evidence for segmentation (e.g. large relative orientation) would decouple the coordination between the center and surround. An adaptive response to an increase in evidence for segmentation therefore would reduce the influence of the gain control pool on the central filter activation. From natural image statistics, we demonstrated that the relative contrast and depth cues maintain similar segmentation effects as relative orientation cues: with greater contrast or depth differences, the center and surround are less likely to belong to the same segments in natural scenes (see Appendix, Figure 4.9). In addition, contrast of a center grating could also affect segmentation: it may be easier to distinguish features of a high-contrast stimulus as opposed to a low-contrast stimulus from background. In the condition of low-contrast center and high-contrast surround, the surround stimulus would set stronger influence on the center, even when their orientations are quite different. This may lead to more co-assignment of center and surround units. However, in the opposite condition of high-contrast center and low-contrast surround, the surround is less likely to be included in the same gain pool as the center because the segmentation here is even clearer, and surround stimulus has a weaker influence. These effects could also be expected to arise through inference or learning in natural scenes (Coen-cagli, Dayan, & Schwartz, 2009; Coen-Cagli et al., 2012). To summarize, in our modification of the Schwartz model, the probability of including a surround stimulus within the gain pool of the central detector (co-assignment probability) depends on segmentation cues: relative

orientation, contrast and depth between the center and surround, and contrast of center gratings (see Equation A2).

We allowed eight free parameters to find the best possible fit of this adapted model to our data. Five were used to control divisive normalization, and three to determine co-assignment probabilities: described in detail in Appendix **experiment 1: model**. The model fit shown in Figure 4.4, explains 84% of variance in the data. We also compared several nested models, none of which performed as well as the complete model. For example, when we assumed no contribution of the segmentation cues in the co-assignment probability, the model was inferior, only explained about 66% variance in the data. This indicates importance of relative contrast and depth in deciding co-assignment probability in the model. In all, the amount of contribution that divisive normalization by itself introduced into the central orientation perception would be insufficient to completely explain the data.

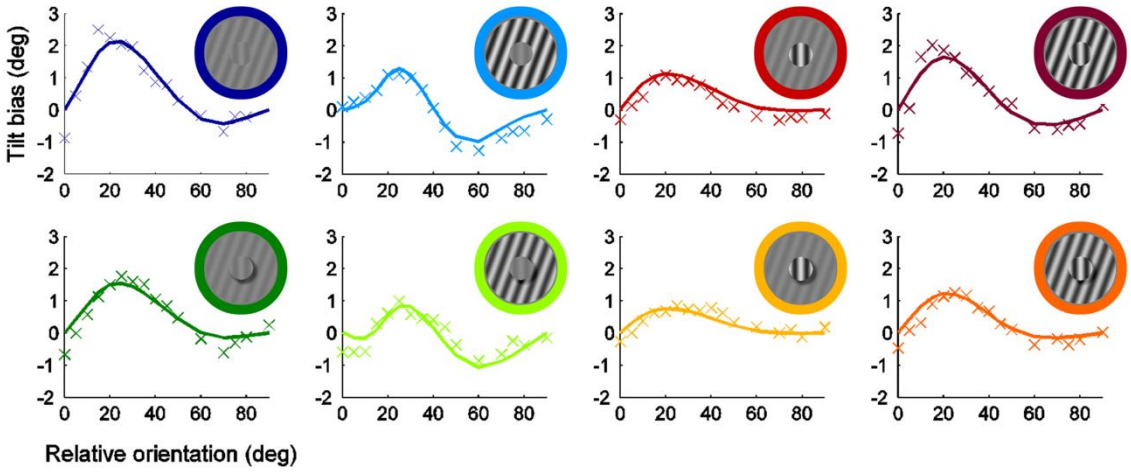

Figure 4.4 Average tilt biases ('x' labels) from eight observers and the least squares fit

(solid lines) of the computational model from eight conditions. When finding the best fit of this model to our data, we fixed four parameters based on the stimuli in each condition, and allowed eight free parameters to fit the 128 data points. When applied to the date with a least squares fit, the model explains 84% of the variance in the data.

4.3 Experiment 2: 2D/3D occluding ring between the center and surround

The segmentation cues in Exp. 1 changed both filtering activations from the center/surround stimuli and the segmentation factor between them, and these manipulations resulted in a reduced tilt repulsion effect. In Exp. 2 we sought to only manipulate the segmentation factor, but leave the initial filtering activation part unchanged, in order to directly examine the effect of segmentation on the tilt repulsion. When relative orientations between the center and surround are small, an annulus covering the boundary between the center and surround may introduce either perceptual segmentation or grouping between the center and surround while maintaining the initial filtering activation. When the annulus is in the same depth plane as the center and surround (2D), it encourages a perceptual interpretation of independence between center and surround, whereas when the annulus is in front of the center and surround in a 3D space, it is more likely to encourage grouping of the center and surround as a common surface. Perceptual grouping through amodal completion has been shown to have effects on perceived transparency (Nakayama, Shimojo, & Ramachandran, 1990) and lightness (Boyaci, Fang, Murray, & Kersten, 2010). We expected a perceptual grouping cue to enhance coordination between the center and surround, thereby increasing the repulsion

effect as in Exp. 1. In Exp. 2 we tested whether this perceptual grouping behind a 3D occluding ring can affect central orientation perception.

4.3.1 Method

Stimuli were as described in Exp. 1 with the center and surround contrast both at 70% and the relative orientations between them at 20 or -20 deg. In Exp. 2 we introduced a 0.2 deg annulus between the center and surround (see Figure 4.5). This annulus was centered on and covered the boundary between the center and surround patch. It was the same luminance as the background and either in the same plane as the center and surround (2D ring) or in front of the center-surround plane in space (3D ring). A stereoscope was used in both conditions. Stimulus duration was 200ms. The boundaries of the annulus and a fixation point were always presented to help maintain fixation. Eight observers (mean age: 27, 5 males) participated in both 2D and 3D ring conditions, and each of their subjective verticals was measured in eight Psi adaptive staircase (Kontsevich & Tyler, 1999) runs, in which four runs were with +20 deg relative orientation and another four with -20 deg (40 trials for each run). When the center-surround relative orientation was 20 deg, the tilt repulsion would be counterclockwise, whereas for -20 deg, the repulsion would be clockwise. The magnitude of the tilt repulsion bias in 2D/3D ring condition was taken as half the difference between subjective verticals for the 20 deg and -20 deg relative orientation runs to eliminate individual vertical biases.

4.3.2 Results

We first fit a linear mixed model with the four experimental conditions (2D ring with +20/-20 deg and 3D ring with +20/-20 deg relative orientation) as fixed effects, and with different subjects as a random effect. We assigned each condition its own mean in the model and built contrasts of these means to test whether repulsion in the 3D ring condition was stronger than in the 2D condition. Figure 4.5 shows individual and average results in the 2D/3D ring conditions (for simplicity we only present the means of results in 20 deg and -20 deg conditions). Both 2D ($z = 12.2, p < .001$) and 3D ($z = 16.7, p < .001$) conditions showed a significant tilt repulsion effect. As expected, stronger tilt repulsion effects occurred in the condition with 3D ring than with 2D ring ($z = 3.14, p = .002$).

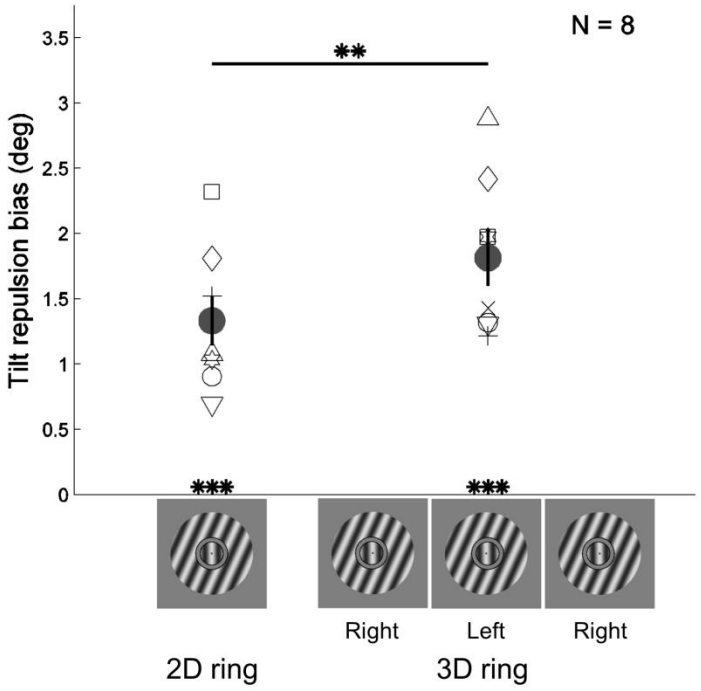

Figure 4.5 Effect of a 2D or 3D occluding ring on the tilt repulsion. The stimuli used in this experiment are shown below the x-axis: 2D and 3D occluding ring (see the stereo image pairs. Left pair for crossed fusion, right one for uncrossed fusion). The legends only illustrate conditions with 20 deg relative orientation. The y-axis shows the tilt repulsion biases, which are the means of results in 20 deg and -20 deg conditions. The gray points show average of eight observers, and error bars show ±1 *SE*. Data for individual observers are shown with smaller icons. $** p < .01$ and $*** p < .001$.

4.4 Experiment 3: Spatial layout of the surround

In the previous two experiments we observed greater tilt biases in conditions with perceptual grouping cues. The spatial layout of surround patches relative to the center

could also be an important factor of perceptual grouping. When surround patches are located collinearly with the central grating, it provides stronger evidence to co-assign the surround and center than when the surround patches flanking the center (Geisler et al., 2001; Sigman et al., 2001). Hence, a stronger repulsion effect was predicted when surround patches were presented at end locations than at flanking locations due to a stronger grouping cue in the former condition. In Exp. 3, we assessed the spatial layout of contextual effects on the perceived central orientation by using stimuli composed of three circular patches: two surround patches were located along different directions to the central grating patch.

4.4.1 Method

Six different layouts, shown in Figure 4.6, were tested: two surround patches positioned vertically (A. end position along the central orientation) or horizontally (B. flanker position), four surround patches as the sum of previous two conditions (C), two surround patches with orientation axes aligned either parallel (E) or perpendicular (F) to the local orientation of the surround, and the sum of the pairs of oblique patches (D). The central patch was 1 deg in diameter as before, and the diameter of the surround patches, which were adjacent to the center, was 1.5 deg. The contrast of center and surround patches was 70%. The boundaries of these patches were slightly blurred using a Gaussian lowpass filter with standard deviation 0.08 deg. In the main experiment (condition A to F), peripheral patches contained 2 cpd gratings; in the control experiment (condition A' to F'), the patch layouts were maintained, but noise with the same spatial frequency was presented instead in order to measure how global orientation of the

surround could influence the center perception (Morgan & Baldassi, 1997; Morgan, Mason, & Baldassi, 2000). Noise patches were generated by filtering white noise in the frequency domain with a Gaussian distribution that was isotropic in orientation, centered about the same spatial frequency as the central grating (2 cpd) with a bandwidth of 0.75 octaves. Stimulus duration was 200 ms. Five observers participated in both main and control experiments. Two additional observers only participated in the main experiment, and two others in the control experiment only. In the main experiment, the relative orientation between the center and surround gratings was 20 or -20 deg as in Exp. 2. The repulsion bias for individual observers in each condition was measured using eight adaptive staircase runs with either 20 deg or -20 deg (four runs for each) relative orientations between the center and surround. The 20 deg and -20 deg conditions should both induce repulsion, and running under these two conditions was to eliminate the individual-dependent subjective vertical offset. The magnitude of the bias plotted in Figure 4.6 is the mean of subjective vertical for the 20 and -20 deg surround orientations. In the control experiments, all patch layouts matched those in the main experiments. However, since there is no orientation information in the surround patches in the control conditions, we characterized bias as leftward (counterclockwise) rather than repulsive or attractive. The y-axis of Figure 4.6 therefore plots the magnitude of the bias associated with sample stimuli along the x-axis rather than the signed repulsion and attraction shown for previous experiments.

4.4.2 Results

As in Exp. 2, we first fit a linear mixed-effect model with each condition having its own mean in the model as a fixed effect, plus a random effect from the subjects. Then we built contrasts of these means to test differences of interest among conditions. The central perception of tilt was significantly biased by peripheral patches with gratings in all conditions (Figure 4.6). For the noise patches, only conditions with oblique patches (E', F') showed tilt biases. The positions of surround noise patches provide global orientation information and induce repulsion and attraction effects on the central gratings in condition E' and F', respectively. We also noticed that, in this case, the magnitude of attraction to the surround global orientation in condition F' was stronger than the magnitude of repulsion in E' ($z = 8.05, p < .001$). However, by adding these oblique surround patches together (as in condition D'), the global orientation information was disrupted, causing the perceptual bias to vanish. The surround area is another important factor that contributes to the magnitude of the contextual modulation, as suggested by Petrov and McKee (2006). This is also true on the central perception of tilt. After taking account the effects of global orientation measured by the control conditions (e.g., A-A'), the four-patch surround caused stronger tilt repulsion bias than the two-patch surround (e.g., C-C' > A-A', with $z = 3.70, p < .001$), and the net effects of two 2-patch conditions were not significantly different from the condition with four patches (e.g., C ≈ A+B, with $z = 1.34, p = .180$).

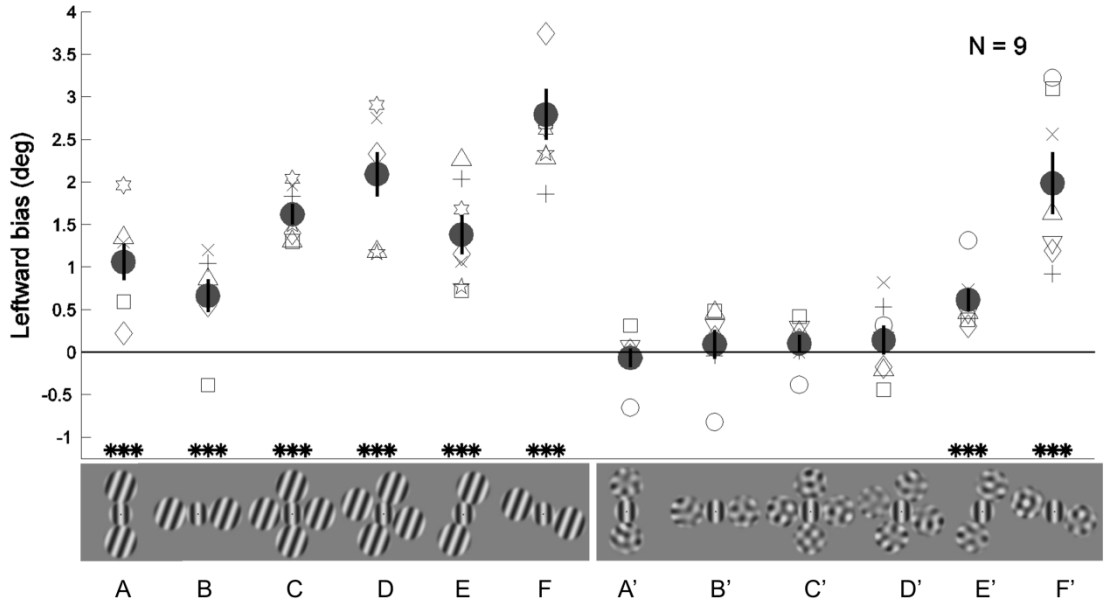

Figure 4.6 Dependence of tilt illusion on spatial layout around the center. Left (A to F): example stimuli and results from the main experiment; right (A' to F'): example stimuli and results from the control experiment. The magnitude of leftward bias plotted along the y-axis is the mean of subjective vertical for 20 deg and -20 deg relative orientations. Example stimuli are shown along the x-axis, and the legends only illustrate conditions with 20 deg relative orientation. Biases are significantly different from 0 in all conditions except A', B', C' and D' ($*** p < .001$). In condition E' and F', the positions of noise patches provide global orientation cues in the surround, and induce the tilt bias in the center.

The original question in this experiment was to assess the spatial layout of the surround induced orientation bias. We used noise patches in the control experiment (same location of surround patches as in the main experiment, but presented with band-pass noise instead of gratings) to discount the influence of global orientation from patch

positions in the main experiment. After subtracting the control effects (see Figure 4.7A), multiple comparison was performed among conditions A-A', B-B', E-E' and F-F'. Figure 4.7B shows estimated biases in these conditions based on measurements from the nine observers: the biases are not equal across four conditions. In particular, the bias in the condition with horizontally presented patches (B-B') is very weak compared to the others. Condition E-E' shows an unexpectedly large repulsion bias, which may result from nonlinear effects that cannot be fully discounted by the control; or, since the standard error in this condition is large, it may not bias center orientation perception in a consistent manner. Only the difference between condition A-A' and B-B', that is, end positions versus flanking positions along the central orientation, is significant: the adjusted p-value is .0344 (with *Bonferroni* correction). Surround patches located along the end positions of a central grating induce a stronger repulsion bias than those on the flanking locations after the global orientation of patch location is accounted for.

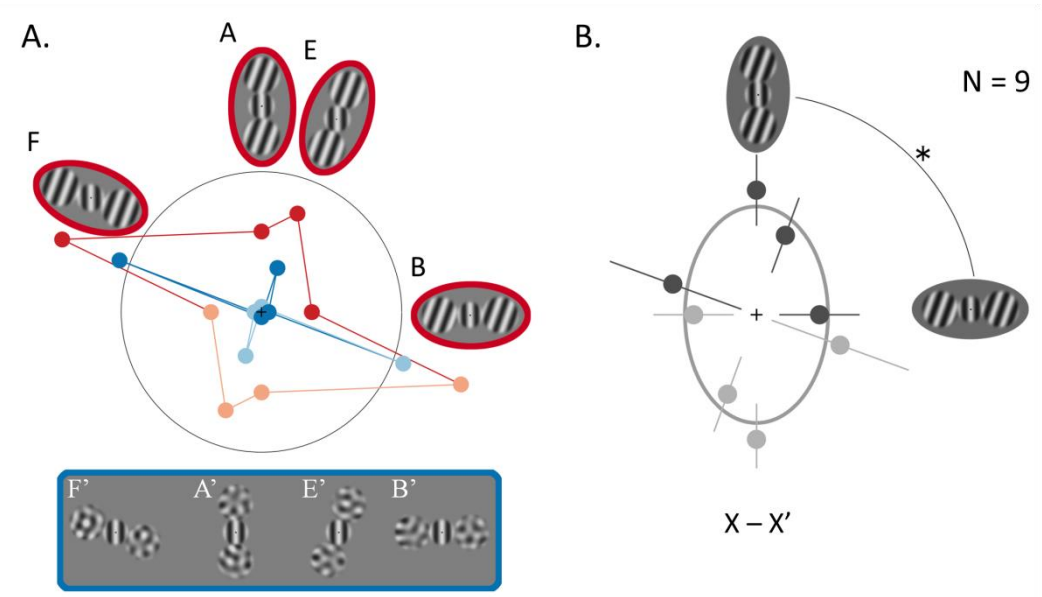

Figure 4.7 Tilt biases induced by different spatial layouts of surround patches. A. Tilt biases in the main (red) and control (blue) experiments. Example stimuli are shown by the icons, the same as the condition A('), B('), E(') and F(') in Figure 4.6. The biases are plotted in polar coordinates. Polar angle indicates global orientation of these patches. Radius shows the strength of leftward bias. Radius of the gray circle shows the bias with a whole annulus surround grating estimated from results in Exp. 1 (1.85 deg). B. The tilt repulsion biases from the same four conditions after subtracting the control effects (condition X-X'). Black dots show average results from four conditions we actually tested, and gray dots are projected for visualization. The biases are not equal across all four conditions. The gray curve is the best fit ellipse ($a = 0.980, b = 0.650$) of these tilt biases across spatial layouts of surround patches. When surround patches were placed along the end positions of the central grating (A-A'), the bias differs significantly

($p = 0.0344$, corrected) from the condition with the horizontally presented surround patches (B-B'). Error bars show ±1 *SE*.

4.5 Discussion

In order to understand how scene segmentation cues affect the tilt illusion, we performed three experiments in which depth, contrast or surround geometry were manipulated. We first measured the effect of two sources of segmentation information, center-surround relative contrast and stereo disparity, on the strength of the tilt illusion in human observers (Exp. 1). Both segmentation cues perceptually decouple the center and surround and reduce the tilt effect. Our results on the relative depth are consistent with Durant and Clifford (2006) but not Sakai and Hirai (2002) or Westhhmer (1990). Sakai et al. used two bars forming an X-shape in their psychophysics and showed that the tilt effect was almost the same regardless of variations in stereo disparity between the target and contextual bars. Westhhmer (1990) also used short lines. It is possible that the depth effect here is induced by the difference in surface assignment, not stereo disparity per se (Huang et al., 2012). Gratings used in our study and Durant et al.'s provide surface segmentation information, while stimuli from Sakai et al. and Westheimer rely more on local stereo disparity.

We also observed that the conditions with a low-contrast center but high-contrast surround show much stronger attraction. The low contrast reduces visibility of the center, which may require increasingly large amount of information from the surround in order to get the central orientation (Mareschal, Morgan, & Solomon, 2010), which potentially increases assimilation of central features to the surround (i.e., attraction). Or

in this scenario, surround effects are relatively stronger when the center is weakly driven (Coen-Cagli et al., 2012), which may lead to more co-assignment of surround units to the gain pool even when the center and surround orientations are quite different, thus causes stronger attraction. However, in the case of high contrast in the center and low contrast in the surround, the surround is less likely to be grouped and the center is easier to be perceived, thus the orientation biases could be reduced. To summarize, the contrast of center grating also matters in the perceived central orientation.

In Exp. 1 we adapted the Schwartz model (Schwartz et al., 2009), which combines both divisive normalization and segmentation factors to fit the psychophysical results from our eight experimental conditions. A key feature of the model is its consideration of perceptual segmentation cues that determine the co-assignment probability of surround stimuli within the gain pool of a central detector. Cues such as center-surround relative orientation, contrast and stereo disparity influence this co-assignment probability, which is crucial in explaining the data we have. For example, stronger input in the surround than in the center (e.g. the condition with high-contrast surround but low-contrast center) can direct the tilt effect toward the attraction, including the repulsion decrease and the attraction increase. These results cannot be predicted well by a traditional divisive gain control model.

As shown in Figure 4.8, a surround grating with greater contrast (light blue) induces a stronger effect on the gain pool than the condition with the same low-contrast center and surround (dark blue), which successfully predicts more reduction of the overall population response in the center of the former condition (Carandini & Heeger,

2012). The stronger gain effect could also push the population codes of the perceived orientation farther away from the real center orientation (a stronger repulsion shown in Figure 4.8). However, this is inconsistent with our observation that the condition with a higher contrast surround shows much weaker repulsion than the condition with low-contrast center and surround (see the dark blue and light blue dots in Figure 4.3). Introducing the segmentation factor can better account for this effect: when the center and surround orientations are similar, contrast difference between the center and surround decreases the co-assignment probability, makes the visual system less likely to assign the center and surround into the same gain pool, and reduces the repulsive bias. We separately manipulated this segmentation factor in Exp. 2 by keeping the same center and surround contrast and orientation but changing the co-assignment probabilities, and we did see variations of tilt associated with this factor.

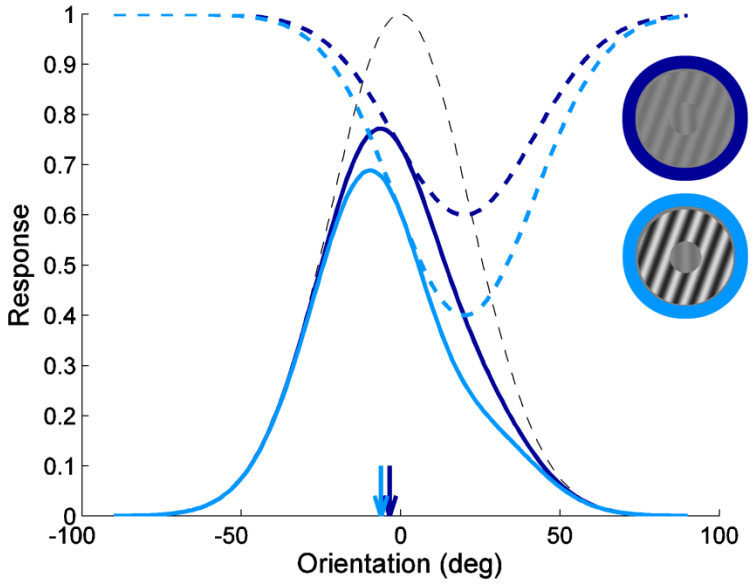

Figure 4.8 Predicted perceptual tilt biases from a divisive gain control model (McDonald, Seymour, Schira, Spehar, & Clifford, 2009; Solomon et al., 2004). The gray dashed line

shows hypothetical population neural response to a vertical central grating without any contextual stimulus. The bluish dashed lines show, when the surround grating is oriented 20 deg relative to the center, its effect on the gain of the neurons responding to the center. The dark blue represents the condition with a low-contrast center and low-contrast surround, whereas the light blue shows the condition with the same center but high-contrast surround. The solid lines are predicted neural population responses to the central gratings surrounded by 20 deg grating with different contrast as shown in icons. The predictions are calculated by multiplying the response to the center-only condition (gray dashed line) by the bluish dashed lines. The condition with higher surround contrast elicits stronger tilt repulsion (farther away from the vertical), which is inconsistent with our results from Exp. 1.

In Exp. 2, a 2D occluding ring (as a gap) spatially separates the center and surround and reduces the tilt effect, which are also shown in Wenderoth and Johnstone (1988) and Durant and Clifford (2006); while a 3D occluding ring encourages center and surround to be grouped as the same surface, and thus increases the co-assignment probability, resulting in stronger repulsion. Functionally, with extra perceptual grouping cues, the visual system may tend to increase the importance of inferring a "hidden" orientation-texture mismatch, and it leads to a stronger bias in this case; whereas a clear 2D gap would make it unnecessary to overemphasize the discrepancy between the center and surround (Durant & Clifford, 2006).

In Exp. 3, we observed that the tilt repulsion effect is strongest along the ends of the stimulus as defined by the axis of central orientation (collinear), which may be because the high edge co-occurrence rate along that location makes observers more likely

to group those surround gratings with the center. The statistics of natural scenes suggests greater orientation dependence between collinear elements than parallel elements, which reflects the predominance of elongated boundaries in the environment (Geisler & Perry, 2009; Geisler et al., 2001; Sigman et al., 2001). To adapt to these statistics in natural scenes, observers would show bigger co-assignment probability in the collinear condition, which follows the local grouping function proposed by Geisler and colleagues (2001). This collinear grouping could also be explained by a generalized form of divisive normalization model using learned or inferred covariance matrices from natural scenes (Coen-cagli et al., 2009, 2012), which successfully predicts a higher co-assignment probability for the collinear condition. When the center and surround are more coordinated, the system tends to exaggerate mismatches of the target from its context, which is represented as a stronger repulsion bias. This process helps to emphasize the discrepancies of actual inputs from the prior belief of the system.

On one hand, the visual system sets a higher co-assignment probability for inputs which are more likely to be the same, and this ensures coding efficiency. This is generally consistent with Cavanaugh et al. (2002b) and Z. Li (2002): neurons respond less to uniform stimuli and more to targets that are distinct from their context. On the other hand, with a high co-assignment probability, the system tends to exaggerate mismatches between elements, and to be more sensitive to the potential mispredictions. Using a set of similar stimuli, Mareschal et al. (2001) found that in the "collinear" condition, orientation discrimination thresholds were significantly bigger than in the flanking condition. It is possible that the exaggeration of orientation bias along the end

positions sacrifices the system's sensitivity to actual angles, but makes it focus more on discrepancies between the center and surround. In all, our results demonstrate that the tilt illusion is affected by spatial layouts of the surround mask, and this spatial anisotropy of the contextual effects may be related to the statistical features of edge co-occurrence relative to the center.

However, our results in Exp. 3 are inconsistent with Kapadia et al. (2000), in which they presented three small line segments (each about 0.13 deg in length) in the fovea with viewing distance of 6 meters. They observed stronger repulsion effects with lateral flankers than with collinear flankers when the relative orientation between the target and flanks was 20 deg. Also, a recent paper by Mareschal and Clifford (2013) reported that surrounding locations equally contributed to contextual effects which did not show any collinear structure. Different patterns in the results may be due to different stimuli used in these experiments, which induce different segmentation between the center and surround.

To summarize, we observed that the tilt repulsion biases increase as it becomes more difficult to perceptually separate the center and surround. Similarity or co-assignment between the center and surround stimuli increases the repulsive shift between the perceived center and surround orientations, which is apparently against our intuition that if the center and surround stimuli become more similar, we would expect our perception of them to be more similar. However, the visual system amplifies the discrepancy among environmental cues which have other evidence of common coordination. This may actually play an important role in contour detection and figure-

ground segmentation. For an example of breaking camouflage, multiple sources of information (luminance, contrast, or color) may seem to say that it is only a bunch of dead leaves or uninteresting bark, but subtle clues (e.g., differently oriented boundaries) tell us that a butterfly is embedded in the background. The visual system must search for and detect camouflaged objects, while at the same time striving for efficiency. Therefore, interactions between the center and surround should not only achieve coding efficiency, but also control the importance of inferring a potential feature mismatch. This high sensitivity to feature contrast between the target and its context, especially in situations that seem to have a common source, could essentially benefit our visual search performance.

The effect of segmentation on the tilt illusion induced by relative orientations and other sources, such as relative contrast, disparity depth, and geometric features, may have different mechanisms. If we assume that the effect of adding relative contrast is the same as increasing the relative orientation, then the effect we see at 20 deg relative orientation with relative contrast should be equal to the condition, say, at 30 deg relative orientation without relative contrast. Therefore, the tilt bias curve as a function of relative orientation should shift toward the left, when relative contrast is introduced between the center and surround. However, in Exp. 1, with relative contrast or disparity depth, the tilt bias curves tend to be right-shifted instead. This suggests that these relative cues may influence the orientation perceptual bias through different mechanisms.

Our results agree with former work by van der Smagt et al. (2005) in which contrast and orientation segmentation cues were used for investigating the role of V1

cells in surface segregation: though a surround of either the same orientation or the same contrast has a suppressive effect on the response to the central stimulus, the authors found that combining the two cues had no greater effect than one on its own. Another similar finding is by Clifford, Spehar, Solomon, Martin, and Zaidi (2003) in which a segmentation cue, color, was used: they found that the tilt repulsion was greater when the center and surround were the same color. This pattern is also true in Durant and Clifford (2006): when the center and surround are perceptually segregated by asynchronous presentation or spatial cues other than orientation, the tilt repulsion effect on the center is reduced. Just as suggested by the authors, if the mechanism underlying the tilt illusion tends to segment surfaces by emphasizing the difference in orientation, when surfaces are already segmented by other cues, the exaggerate changes of orientation are not that crucial (Durant & Clifford, 2006). On the other hand, if those cues aid perceptual grouping between the center and surround, the tendency of emphasizing the orientation difference would be enlarged.

In order to demonstrate different mechanisms or explore the level of segmentation information processing, a backward noise masking of the surround (Clifford & Harris, 2005) or a rapid reverse-correlation method Mareschal and Clifford (2012) could be useful. In a recent paper by Mareschal and Clifford (2012), the authors suggested that a single mechanism operating in the early stages of visual processing (before conscious perception of the surround) could account for both the tilt repulsion and attraction. They used a reverse-correlation technique, in which the surround orientation was changed every 12 ms making it invisible to observers. They found that both the tilt repulsion and

attraction occurred over a similar time course, which suggested that it may not be necessary to invoke a separate, higher-level mechanism. It will be interesting to see whether the effect of the perceptual grouping/segmentation cues used in our experiments persists when the surround orientation is not consciously perceived. This reverse correlation paradigm may help to entangle the levels of processing involved with different center-surround perceptual grouping/segmentation cues.

4.6 Conclusion

In conclusion, our results from three experiments demonstrate that center-surround relative contrast, relative disparity depth, and geometric features, such as occlusions and colinearity, can affect the strength of perceptual orientation bias in the center. In general, a stronger perceptual grouping cue between the center and surround enhances tilt repulsion biases, whereas a segmentation cue reduces the effect. Functionally, this may increase the sensitivity of our visual system to feature discrepancies, especially in an environment rich in similarities, and this may play an important role in visual search and detection.

4.7 Appendix

Table A1.

Experiment 1 statistics results

Factor	Maximum repulsion	Relative orientation at maximum repulsion	Maximum attraction	Relative orientation at maximum attraction	Cross-over orientation

Contrast	$F_{(1,33)} = 58.7$ $p < .001$	$F_{(1,33)} = 2.05$ $p = .162$	$F_{(1,33)} = 4.94$ $p = .033$	$F_{(1,33)} = 5.87$ $p = .021$	$F_{(1,33)} = 1.56$ $p = .221$
Depth	$F_{(1,33)} = 21.3$ $p < .001$	$F_{(1,33)} = 24.4$ $p < .001$	$F_{(1,33)} = 6.21$ $p = .018$	$F_{(1,33)} = 4.42$ $p = .043$	$F_{(1,33)} = 7.84$ $p = .008$
Central contrast	$F_{(1,33)} = 6.25$ $p = .018$	$F_{(1,33)} = 0.0766$ $p = .784$	$F_{(1,33)} = 31.0$ $p < .001$	$F_{(1,33)} = 12.1$ $p = .001$	$F_{(1,33)} = 9.16$ $p = .005$
Contrast : depth	$F_{(1,33)} = 2.49$ $p = .124$	$F_{(1,33)} = 0.603$ $p = .443$	$F_{(1,33)} = 3.99$ $p = .054$	$F_{(1,33)} = 0.322$ $p = .575$	$F_{(1,33)} = 1.43$ $p = .241$
Contrast : central contrast	$F_{(1,33)} = 1.89$ $p = .179$	$F_{(1,33)} = 0.687$ $p = .413$	$F_{(1,33)} = 22.2$ $p < .001$	$F_{(1,33)} = 15.6$ $p < .001$	$F_{(1,33)} = 23.2$ $p < .001$
Depth : central contrast	$F_{(1,33)} = 0.419$ $p = .522$	$F_{(1,33)} = 0.245$ $p = .624$	$F_{(1,33)} = 0.272$ $p = .605$	$F_{(1,33)} = 0.178$ $p = .676$	$F_{(1,33)} = 1.90$ $p = .177$

Note: A:B indicates interactions between the factor A and B.

Experiment 1: model.

Influenced by the gain control pool, the estimate of the normalized neural response associated with the central detector is:

$$E(g_{ci}|l_{ci},l_{si}) = \frac{l_{ci}}{\sqrt{l}} \cdot \frac{\mathrm{B}(\frac{n}{2}-\frac{1}{2},l)}{\mathrm{B}(\frac{n}{2}-1,l)} \qquad \text{(A1)}$$

where $l_{ci}$is the filtering response of the central detector tuned to a particular orientation ϕ_{ci} (preferred orientation), $l_{ci} = C_c^{C_{exp}} \cdot \exp(-(\phi_{ci} - \theta_c)^2/2\varpi_c^2)$, when the orientation of center stimulus is θ_c, and the contrast is C_c; $l_{si}$is the response to the surround stimuli oriented at θ_s with contrast of C_s, $l_{si} = C_s^{C_{exp}} \cdot \exp(-(\phi_{si} - \theta_s)^2/2\varpi_s^2)$, where C_{exp} controls the contribution of contrast in filter responses; and the gain control pool for detector *i* is set by center and surround filter activations with the same orientation

preference, the divisive term is $l = \sqrt{l_{ci}^2 + (n-1)l_{si}^2 + k}$, where n describes the strength of surround influence on the gain pool (one can think n is related to surround size relative to center), and *k* is an additive constant; B(·) stands for a modified Bessel function of the second kind.

Another key component in the model is the segmentation factor. Modified from the Schwartz model, the probability of including a surround stimulus with the orientation θ_s within the gain pool of the central detector (co-assignment probability) depends on all possible segmentation cues (orientation, contrast and depth):

$$p = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(\frac{\phi_{ci}-\theta_s}{\sigma})^2} \qquad \text{(A2)},$$

where $\sigma = \lambda e^{C_c + C_{weight}\cdot(C_c - C_s) + D_{weight}\cdot(D_c - D_s)}$ gives the steepness of the co-assignment selection, which is determined by central contrast C_c, relative contrast $(C_c - C_s)$ and relative depth $(D_c - D_s)$. When there is no contrast difference between the center and surround, we have $e^{C_{weight}\cdot(C_c - C_s)} = 1$, which shows no effect on co-assignment; and in those conditions with contrast difference, the relative contrast is positively weighted by C_{weight}. Based on results in Exp. 1, the center contrast could also affect co-assignment. C_c, therefore, is another component in determining the steepness. $e^{C_c + C_{weight}\cdot(C_c - C_s)}$ is the same as $e^{aC_c + (1-a)C_s}$, where $a = 1 + C_{weight}$ and C_{weight} is a positive number. The higher the center contrast is, the shallower the slope of co-assignment probability is, and the smaller the peak co-assignment is; whereas the higher the surround contrast is, the steeper the co-assignment slope is, and the greater the peak co-assignment is. This is consistent with our expectation: when the surround contrast is constant, a higher contrast

center would decrease the maximum probability of including a surround stimulus within the gain pool of the central detector; while when the center contrast is constant, a higher contrast surround would increase the maximum co-assignment probability. As for relative depth, when there is no depth difference between the center and surround, we have $e^{D_{weight}\cdot(D_c-D_s)}=1$, which has no effect on co-assignment; and in those conditions with depth difference (in our experimental conditions assuming $D_c \geq D_s$), the relative depth is positively weighted by D_{weight}. The greater the relative depth is, the smaller the peak co-assignment is. In other words, depth difference would decrease the co-assignment probability.

If the surround stimulus is not taken as being part of the same gain pool as the center detector, then the detector would take into account only the center stimulus $E(g_{ci}|l_{ci})$. The net response is weighted by the co-assignment probability and is given by:

$$pE(g_{ci}|l_{ci},l_{si})+(1-p)E(g_{ci}|l_{ci}) \quad \text{(A3)}.$$

Then through standard population decoding (Georgopoulos, Schwartz, & Kettner, 1986),

$$r=\frac{1}{2}angle(\sum_i g_i\vec{u}(2\phi_{ci})) \qquad \text{(A4)},$$

we obtain the perceived central orientation r.

When finding the best fit of this model to our data, we fixed C_c, C_s, D_c and D_s based on the stimuli in each condition ($C_c=0.1\ or\ 0.7$, $C_s=0.1\ or\ 0.7$, $D_c=0$ and $D_s=0\ or\ 2$), and allowed eight free parameters: n, k, center tuning width ω_c, surround tuning width ω_s and C_{exp}, respectively; plus C_{weight}, D_{weight} and λ, respectively, when calculating co-assignment probability. Average data from all eight conditions are

summarized in Figure 4.4. Sample size is 128 (16 points in each of 8 conditions). When applied with the least squares fit to the model, we obtained optimal parameters:

$n = 5.4, k = 0.25, \omega_c = 17\ deg, \omega_s = 10\ deg, C_{exp} = -0.60, C_{weight} = 1.7, D_{weight} = 0.19\ and\ \lambda = 44\ deg$

. Fit results are shown as solid lines in Figure 4.4, and it explains 84% of the variance in the data. In an attempt at parsimony, we obtained fits to several nested models. The small-sample-size corrected Akaike Information Criterion (AICc) was used to evaluate these models. A smaller AICc indicates a more efficient fitting. In one of the nested models, we forced the probabilities not associated with the relative contrast and depth, that is, set $C_{weight} = 0$ and $D_{weight} = 0$. The model was inferior (AICc = -223) to the original model (AICc = -317), and only explained about 66% variance in the data (versus 84%), indicating importance of relative contrast and depth in deciding co-assignment probability in the model. In another nested model, the effect of center contrast on the co-assignment was eliminated, that is, the condition with high-contrast center and low-contrast surround and the condition with low-contrast center but high-contrast surround had the same co-assignment probability. This fit explained 77% variance in the data with AICc = -268.

Experiment 1: model – natural image statistics.

In order to demonstrate that contrast and depth cues maintain similar segmentation effects as the orientation cue, we measured the joint conditional distribution of the contrast or disparity depth in the center, given the contrast or disparity in the surround. The center was defined as a 9×9 pixel square patch, and the surround as one of four possible edge-adjacent neighboring 9×9 pixel patches. When a continuous

contour longer than eight pixels was detected within a 6×9 pixel patch centered on the boundary between the center and surround patches, the patches were classified as across boundaries, otherwise, they were said to be within boundaries. Contrast within a given patch was measured based on Michelson contrast. Disparity depth was the mean of the depth value for all pixels in the patch. Figure 4.9 shows that the correlations of contrast and disparity depth between the center and surround patches are reduced across boundaries from pictures in the Berkeley database (Martin, Fowlkes, Tal, & Malik, 2001) and the stereo depth database (Scharstein & Pal, 2007) respectively, suggesting that the center and surround tend to be more separated due to contrast and depth cues. Figure 4.9D and 4.9H further show that the probability of central and surround patches belonging to the same surface or object (within boundaries) decreases as the center-surround contrast (or depth) difference increases.

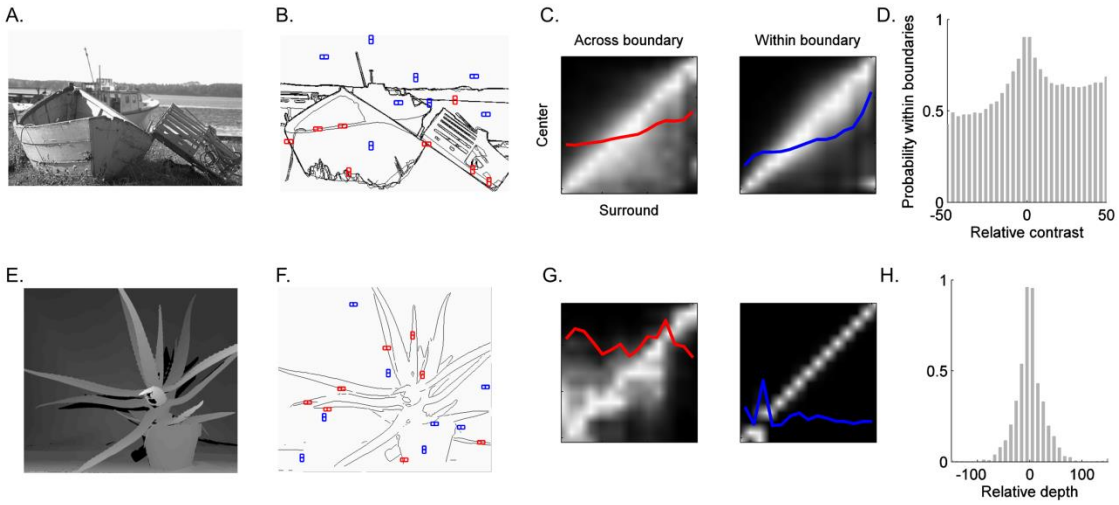

Figure 4.9 Statistical dependencies in term of contrast (A-D) and depth (E-H) information in natural images within (blue) and across (red) boundaries. The statistics for each condition were collected over 30000 random samples from each of 9 images. A.

Example image from the Berkeley database (Martin et al., 2001) including hand-labeled segmentation boundaries (as shown in B). C. The joint statistics between center and surround patches when they belong to different segments (left, red) or the same segment (right, blue). The plots show the joint conditional statistics of the contrast in the center, given the contrast in the surround. Contrast within a given patch was measured based on Michelson contrast. The bottom left shows the count of the center contrast being 0 when the surround contrast was 0. Intensity is proportional to the counts, but each column is independently re-scaled to show a conditional distribution given a certain surround contrast. The solid lines show the conditional standard deviation. Within boundaries, the center and surround patches tend to have similar contrast. This similarity is greatly reduced across boundaries, suggesting that the center and surround are less likely to be grouped. D. Probabilities of central and surround patches within boundaries as a function of relative contrast between the center and surround. When the center and surround have similar contrast, they tend to belong to the same segment in natural scenes. E. Example image from a stereo depth image database (Scharstein & Pal, 2007), and segmentation boundaries were calculated based on the depth information (as shown in F). G. Joint conditional statistics of the center patch depth, given the depth information of the surround patch. Disparity depth was the mean of the depth value for all pixels in the patch. The bottom left shows the count of the center depth as 0 when the surround depth was 0. The solid lines show the

conditional standard deviation. The blue represents when the center and surround patches belonging to the same segment, and the red represents patches across boundaries. Within boundaries, the center and surround patches tend to have similar stereo depth, whereas this similarity drops across boundaries, implying that the across boundary surround does not provide as much information as it provides to the center when they are belong to the same segment. H. Probabilities of central and surround patches within boundaries as a function of relative stereo depth between the center and surround.

Acknowledgement

This work was funded by R21 NS075525. D.K. was supported by the World Class University program funded by the Ministry of Education, Science and Technology through the National Research Foundation of Korea (R31-10008), NIH grant R01 EY015261, and ONR N000141210883. The authors would also like to thank Damien Mannion, Michael-Paul Schallmo and the reviewers for their comments on the manuscript.

5. Shape distortion with fast repeated presentations

Authors

Cheng Qiu, Cheryl A. Olman, Daniel Kersten

Summary

We observed that fast-paced peripheral presentations of ellipses (including circles) with slightly varying size and aspect ratio result in their apparent change in shape. When observers were asked to sketch the perceived shapes, they all portrayed perception of a closed form consisting of a discrete set of vertices connected by straight edges.

The magnitude of the effect was quantified by the time it took for distortion to appear. Observers were instructed to press a response key when they saw the distortion first begin to appear; the trial ended when a responses was received or after a minute of presentation. We found that: (1) The effect in the opposite direction (perceiving ellipse with polygon presentations) hardly occurs; (2) The increase of shape and size dissimilarity first enhance the distortion, but larger dissimilarities weaken the effect; (3) A larger location shift largely eliminates the shape distortion; (4) The distortion is enhanced with increasing eccentricity, contour spatial frequency, or contrast; (5) The distortion can transfer interocularly; (6) Over a flashing frequency range between 1 and 42 Hz, the magnitude of the effect increases up to about 10 Hz. We also found that the

effect persists when a temporal gap (i.e. a 100ms blank frame) is inserted between stimulus frames.

Taken together, this effect is primarily due to continuous adaptation to local oriented-line or curvature, and global shape integration may enhance a closed polygonal perception. The illusory perception here may reflect an efficient representation of rapidly changing visual inputs, and we showed that this process can happen very early in the visual processing stream.

5.1 Introduction

We observe a shape distortion in which fast-paced peripheral presentations of ellipses with slightly varying size and/or position result in polygonal shape perception. Similar effects were observed with static curvature (Khuu, McGraw, & Badcock, 2002; Watt & Andrews, 1982), inward graduation-flash of circles (Sakurai, 2014a), or in the form of an afterimage (H. Ito, 2012). Watt and Andrews (1982) documented the observation that the ends of large curves appeared “droopy” and a circle seemed to be a smoothed polygon with about ten edges after a short time of fixation. Khuu et al. (2002) reported a peripheral circle would appear polygonal in shape after viewing for a period of time (as induction time). They also found that the induction time reduced linearly with increasing eccentricity of the circular form (either scaled for cortical magnification or fixed at 2 deg radius); they measured the number of sides of the perceived polygon, which was constant at 7 sides for scaled form, but for a fixed-size circle fewer sides were perceived with increasing eccentricity. Also with static stimuli, H. Ito (2012) discovered that observers reported a hexagonal afterimage after viewing a circle for 10 sec;

interestingly, he also found that sometimes observers saw a circular afterimage after viewing a hexagon. Sakurai (2014a) studied this effect in a dynamic mode in which alternations of a circle and its inward gradation were used. A polygonal percept was also reported, and a much shorter viewing period was required: on average, the induction time was about 4 sec when the temporal frequency was 2Hz (the temporal frequency would significantly affect the induction time). In later studies, they further reported that this shape distortion induced by graduation-flash was unidirectional with a much weaker effect of seeing circles by adapting to hexagons (Sakurai, 2014b); and they also demonstrated interocular transfer of this effect (Sakurai & Beaudot, 2015). The distortion we observed could share similar origin as above effects. However, a systematic study of the mechanisms is lacking.

Our fast repeated presentation paradigm was also similar to the one used in the flashed face distortion effect (Tangen, Murphy, & Thompson, 2011), in which eye-aligned faces were presented at a fast pace and when viewed for a short period of time the perceived faces seem to be severely deformed. This can be seen as a special adaptation paradigm. In a traditional adaptation framework, one form is adapted for several seconds to minutes, and another form (similar to the adaptor) then appears as a test frame (a brief interstimulus interval could be inserted in between). The adaptation to the first form tends to bias the perception of the test form toward the opposite direction in a certain feature space, such as size, orientation, motion direction and so on. With repeated presentations in our paradigm, each frame can be treated as both adaptor and test frame. Specifically, each frame among our rapid presentation sequence can be a test frame, and all frames

earlier than this test frame are adaptors: the observed distortions of the test frame are likely due to adaptation to all previous frames. Similarly, an adaptation period was also required in all previously mentioned effects with polygonal percept.

Both low-level (i.e., local) and shape-level (i.e., global) processing mechanisms in visual adaptation have been proposed and intensively studied. The local process is heavily dependent on spatial parameters (spatial overlap of the adaptor and test frame is often required), and has been suggested to be due to selective reduction of the sensitivity of neurons responsive to the adaptor. As a result, the reduced responses cause bias of a tuned feature, such as orientation, size and spatial frequency, in the subsequent test frame (Blakemore & Campbell, 1969; Blakemore, Nachmias, & Sutton, 1970; Köhler & Wallach, 1944). In contrast, the global process happens in figural dimensions independently of spatial parameters. For example, in the shape-contrast effect, a test shape is presented briefly following a rapid priming shape, and the priming shape robustly (across variations in size and position) distorts the perceived test shape so that the two appear more dissimilar (Suzuki & Cavanagh, 1998). A similar effect, the aspect-ratio-based aftereffect, has also been reported (Regan & Hamstra, 1992). These distortions could be explained using a continuous shape coding model (similar as a tuning to orientation, but here in figural dimensions, such as continuous tuning to aspect ratio) – adapting to one shape is likely to bias perception toward the shapes along the opposite tuning direction. This model also predicts bi-directionality of the global process – shapes near either sides of the tuning could be adapted. Indeed, H. Ito (2012) reported perceived afterimages of circle after prolonged viewing of hexagon (in addition to hexagon

afterimage with viewing of circle). Physiologically, neurons in higher levels of the ventral stream, such as inferotemporal cortex, tend to be largely size and position invariant (Brincat & Connor, 2004; M. Ito, Tamura, Fujita, & Tanaka, 1995; Kobatake & Tanaka, 1994), which may server as infrastructures for the global mechanisms. Lastly, the local and global mechanisms in adaptation may not be strictly separate: they could interact to shape the overall perception. As described in the radial frequency shape after-effect (N. D. Anderson, Habak, Wilkinson, & Wilson, 2007), the perceived phase shift is likely to be global since it is invariant to size; but the perceived uncertainty to test shape which modulated by adapting contrast may be due to the local mechanisms. In the present study, we sought to understand the nature of adaptation behind the shape distortion effect we observed.

We first quantitatively measured and evaluated the observed distorted shape. Secondly, in an effort to determine the mechanisms that contribute to this shape distortion, we explored the strength of the effect under a variety of conditions – multiple spatial/temporal variables were manipulated. The distortion could not occur in an opposite direction, that is, no clear ellipse was observed with adapting to a sequence of polygons. Further, it relied on similarities and overlapped retinotopic locations between frames, it could transfer across eyes, and it is sensitive to eccentricity, spatial frequency and contrast. Together, these results suggest that mechanisms responsible for the shape distortion are less likely to be at the global/shape level, not exclusively monocular, but most likely to occur at a local/lower level. A V1 response model was used to simulate the distorted shape, in which adaptation is modeled as a reduction in neural response gain.

Further, we showed that brief presentation (or large retinal eccentricity) inducing uncertainty enhanced the distortion.

5.2 Methods and results

General methods

A total of 20 observers (mean age: 28 years, six males) with normal or corrected-to-normal visual acuity participated in the study. All observers provided informed written consent under an experimental protocol approved by the Institutional Review Board at the University of Minnesota.

Visual stimuli were generated using Matlab (Mathworks, Inc., Natick, MA) in conjunction with the Psychtoolbox (Brainard, 1997; Pelli, 1997). Stimuli were displayed on an LCD monitor (1600 × 1200 pixels, 60 Hz refresh rate, NEC MultiSync LCD 2190 uxi, NEC, Tokyo, Japan) connected to a Mac mini (Apple, Inc., Cupertino, CA). In Experiment 8, in order to control experimental timing more accurately, a CRT monitor (1600 × 1200 Pixels, 85 Hz refresh rate, Mitsubishi Diamond Pro 920, NEC-Mitsubishi Electronics Display of America, Inc., Illinois, USA) was also used. Automatic hardware anti-aliasing of the display was enabled to avoid possible bias introduced by non-smooth boundaries. A chinrest was used to maintain a constant viewing distance of 65 cm away from the monitor. All stimuli were presented binocularly, except that in Experiment 7.2 stereoscope and split screen were used to achieve a dichoptic presentation. At the beginning of Experiment 7.2, the stereoscope was adjusted by aligning two short nonius lines; a square frame and fixation point were remained on the screen during the experiment for stable fusion.

The basic stimuli were consist of a sequence of slightly different ellipses in periphery (as shown in Figure 5.1), and polygons were used in Experiment 2. In a standard test condition, ellipses at 3 °eccentricity with mean minor radius (horizontal) of 1.4 °and mean aspect ratio 4:3 were presented at 4 Hz alternating frequency. All shapes were presented with dark outlines (line width 9.6 minute of arc) on a gray background (-61% Weber contrast). For each ellipse, its major and minor radius were independently drawn from uniform distributions between $(\text{mean major/minor radius}) \pm 2\%(\text{mean major/minor radius})$, where the difference between the largest and the smallest radius, for this example, 4% of the mean radius, was defined as the change amount. These parameters from this standard condition were manipulated in the following experiments.

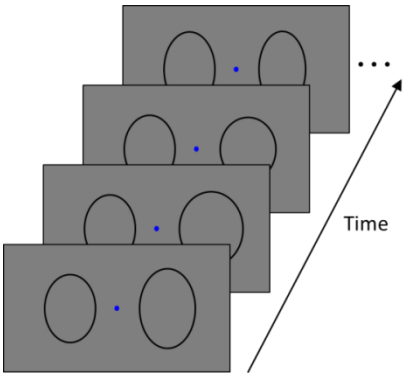

Figure 5.1 Basic stimulus paradigm. Differences across frames were exaggerated for illustration purposes only. Varying ellipses are presented in sequence, and the observers were instructed to fixate at the central blue point.

Observers first went through a practice period. They were well aware that ellipses were presented on both sides of the screen, and they were asked to maintain their fixations at the central blue dot while paying attention to the ellipses and to report what

they saw during a practice phase. The elliptic shape was reported at the beginning, but straight lines, corners, or polygonal shapes were reported after a short period of viewing. Since the stable shape distortion seems to emerge gradually, unlimited viewing time during this practice phase was intended to help establish stable criteria for later response time tests or perceived shape sketch.

The response time/induction time was measured as the time between ellipse sequence onset and observers' response key press. Specifically, observers were instructed to press the Enter key to initiate each new trial, and to press any key when a stable distortion appeared. If no distortion was observed, that is, no response key was recorded, the trial would be automatically terminated in 60 sec. Each trial was followed by a minimum of 20 sec static mask noise to avoid afterimage across conditions. The response time was used to indicate the strength of the distortion, assuming that a stronger effect would be very likely to cause faster response time. The response time to the standard condition (as described above) was measured in 5 repeat trials. In Experiment 1 and 2, observers were asked to sketch the perceived shapes. In subsequent experiments, the response time was recorded, and in Experiment 7 the perceived shapes to cued test frames were collected.

For statistical analysis, most main effects were assessed using Analysis of variance (ANOVA). Interaction effects of the factors were also tested using Analysis of covariance (ANCOVA), which allowed control for the effects of the covariate. Optimized data transformations were applied to stabilize variance if necessary. If the difference between two conditions was of interest, a t-test was conducted.

Experiment 1. Distorted shape sketch

Following the practice, observers were instructed to sketch the perceived shapes from the flashed ellipse presentation, and the sketch was then inserted into the original ellipse sequence for evaluation. This process continued for two to four iterations until a good match was achieved. As mentioned in general methods, a series of ellipses were presented in periphery while observers fixated at the center of the screen. After observers orally described the distortion and the response time was measured, they were instructed to sketch the perceived shape using a Psychtoolbox-based interface. For each iteration, an initial sketch was presented on the screen, and observers were asked to adjust the number of vertices, vertex positions, and edge curvature between the vertices until they thought the sketch matched with the shape they perceived. The initial sketch could be a polygon with the number of vertices that observers selected, or the sketch they made from previous iterations. Observers could freely view the stimulus sequence while they worked on the sketch. They had the option to turn off the black dots on the vertices to compare the sketch with their perceived shape, and the existence of the dots was only as indicator for adjusting areas.

The finished sketch was normalized (so its maximum length/width was identical to the mean major/minor radius of ellipses) and embedded into the sequence of ellipses for evaluating its similarity with the perceived shape: observers were asked to press a response key when they saw their sketch from the ellipse sequence. If the sketch was a good match with the perceive shape, once the distortion happened, it would be hard to accurately identify the sketch from the actual distorted shapes: the hit rate would decrease, whereas the false alarm rate would increase. Five evaluation trials were

conducted. In each trial, 144 ellipses were shown at 4Hz at each side of the screen. About 10% of the ellipses were substituted with the sketch of the distortion, and it could be at either side of the screen. Observers used left and right arrow keys to indicate the presence of the sketch. Each trial lasted 36 sec and followed by a minimum of 20 sec mask noise. For data analysis, each trial was divided into two chunks – before and during the distortion – based on the response time to the distortion measured for individual observer in the practice phase. True and false positives and negatives were separately calculated for these chunks. The Matthews correlation coefficient (Matthews, 1975) was then obtained from above values based on the equation (*), as a balanced measure of detection performance:

$$\mathrm{MCC} = \frac{\mathrm{TP}\times\mathrm{TN}-\mathrm{FP}\times\mathrm{FN}}{\sqrt{(\mathrm{TP}+\mathrm{FP})(\mathrm{TP}+\mathrm{FN})(\mathrm{TN}+\mathrm{FP})(\mathrm{TN}+\mathrm{FN})}} \quad (*)$$

where TP stands for true positives (hits), TN stands for true negatives (correct rejection), FP is false positives (false alarm), and FN is false negatives (miss). A smaller value indicates a better merge between the sketch and the distorted shape.

A good sketch should be easily identified during initial frames (before distortion), and start to merge with the distorted shapes once the distortion occurred. If the sketch was too similar to an ellipse, no such distinction would show, though the sketch did merge well into the sequence. Therefore, for each observer, a best sketch was selected from the iteration that showed the largest difference between the MCC measurements before and during the distortion. The results are shown in Figure 5.2. Though large variations across observers' sketches were observed, the vertices near the shoulder positions were salient for all observers.

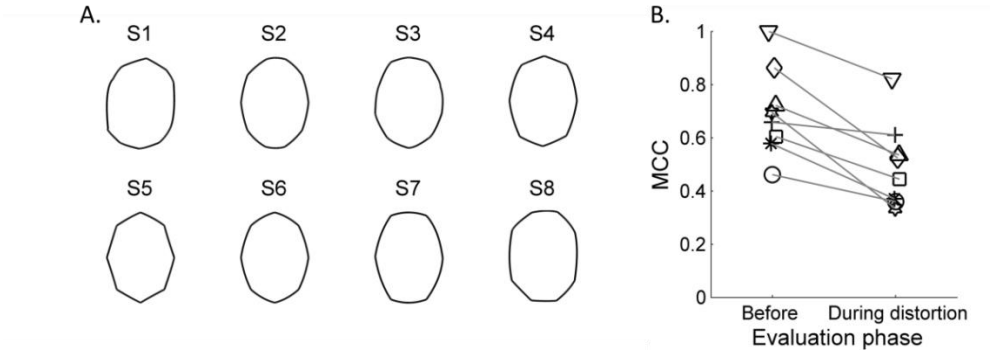

Figure 5.2. Distortion sketch using a computer-based interface. A. The best sketch made by individual observers. B. The Matthews correlation coefficient (MCC) acquired before and during the distortion for the iteration with the best sketch for individual observers (shown in different icons). All selected sketches could be identified more easily before than during the distortion phase.

Experiment 2. Shape distortion with flashed polygons

As mentioned in the introduction, if the effect is due to the global mechanisms, invariant to size or location (non-retinotopic) should be expected, and the distortion should be bi-directional as well. In the case presented here, we can imagine the existence of a continuous shape tuning to the number of sides, e.g., a tuning from triangle, square, hexagon, to decagon, and to circle (an infinite number of sides). When the visual system adapted to round shapes, a biased polygonal percept would emerge. To test this shape-tuning hypothesis, four experiments were conducted. In Experiment 2 we tested whether the effect would occur in the opposite direction, that is, whether a sequence of polygonal presentation could induce circular percepts. Experiment 3, 4 and 5 explored the dependency on stimulus similarity and location.

In Experiment 2, slightly varying polygons with eight edges (symmetric about the horizontal and vertical axis) were presented in the periphery at 4Hz (Figure 5.3A). The

number of sides was decided based on the observers' sketches in Experiment 1. The change amount of the polygon width/height was 4% of the polygons' mean width/height. Observers were instructed to first describe the perceived shape, and then make a hand-drawn sketch with paper and pencil. The sketch was then scanned and converted to binary image as shown in Figure 5.3B. Six observers were tested, all reported seeing bump at the corner, among them one observer experienced almost elliptic shape perception (S5 in Figure 5.3), and another observer reported inward curvature of straight edges (S6).

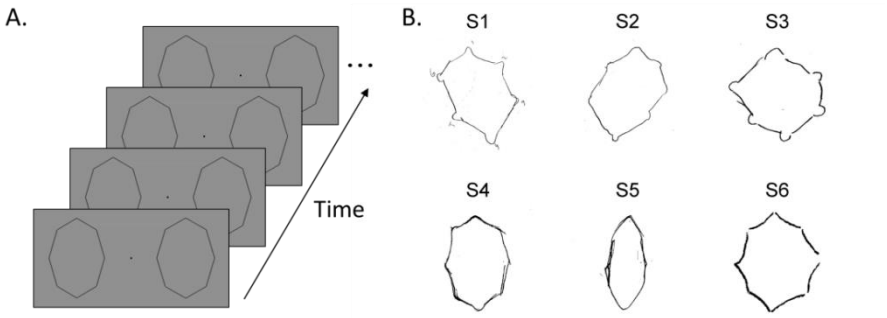

Figure 5.3 The stimulus sequence and the hand-drawn sketches. A. An example of the stimulus sequence. B. The hand-drawn sketches from five observers (converted to binary image).

These sketches suggest that the distortions between ellipses and polygons are not simply bi-directional. Though one out of six observers reported seeing round shapes sometimes, the perception of exaggerated corners is hard to ignore. Our results are consistent with Sakurai (2014b), in which only about 12% octagons were reported as being distorted into circles, whereas 79% distortions from flashed circles to polygons were reported with a 10 sec presentation trial. Even in H. Ito (2012), two clear peaks (both hexagon and circle) could be seen from their results, when the observers were

instructed to report the perceived afterimage shape after adapting to hexagons. The distortion observed in Experiment 2 may indicate adaptions between rectilinear/angular and curved contour segments. One possible explanation for the effect is that local straight edges tilt away from their original orientations with adaptation (as in the tilt aftereffect), which results in discontinuity (gaps) between the edges. In order to maintain the closed form or to bridge the gap, the part near the endpoints could be locally curved. The perceived curvature from straight lines has been reported in a static visual illusion, the lemon illusion (Strother, Killebrew, & Caplovitz, 2015), in which a similar exaggeration of the lemon bulges and illusory concavities were perceived. This may also reflect the role of contour discontinuities in shape perception. Further investigation would be necessary to explore the nature of the effect here. To summarize, we think local features of the shape do matter in this distortion induced by fast repeated presentations, and the distortion does not equally occur along both directions. The perceived distortion from ellipses to polygons is robust, but not for perceiving ellipses after adapting to polygons.

Experiment 3. Similarity between frames

For the proposed shape-tuning hypothesis, no exact circular shape was identified. We have noticed that the distortion also works with repeated circles instead of ellipses. In this experiment, we explored the importance of shape similarity in perceiving the polygonal distortion, specifically, to answer the question whether circular shapes in general are sufficient to induce the distortion or circular shapes with similar features are required. If the distortions are tuned to specific shape features, large dissimilarity would reduce the distortion; whereas if the distortions are less selective to exact shape features, a higher-level mechanism would be expected. The test geometry was the same as the

standard condition (ellipses at 3 °eccentricity with 1.4 °mean minor radius), and the response time from 7 levels of change amount (0, 2, 4, 8, 16, 32, and 64%, as defined in *General methods*) were measured. Since for each frame the width and height were independently drawn from uniform distributions with predefined intervals, both size and aspect ratio of ellipses would change from frame to frame, and with a larger amount of change, the shapes would also be retinotopically less overlapped.

Figure 5.4 shows the results. The average time to see the distortion for the static (0%) condition was 6.6 sec. A U-shaped pattern was observed: when the change was small or excessively large, it took longer for observers to perceive the distortion. Trend analyses revealed that there was indeed a quadratic component to the relationship, $F(1,174) = 20.38, p < 0.001$, with the lowest average response time 4.3 sec for the 8% change condition. The average response time for the static condition was 6.9 sec, and it was significantly longer than the response time when 8% changes were applied ($p < 0.05$, multiple comparison with Tukey's HSD test). The static condition was also similar to the one used in Khuu et al. (2002) and thus the change is not necessary to induce the polygonal distortion. Additionally, when the amount of change was 64%, the average response time was 8.83 sec, which was significantly longer than all other conditions ($p < 0.05$ in each case, Tukey's HSD test). Taken together, the right amount of variation between frames enhanced the distortion, but the excessive change did reduce the distortion, which suggests the distortions largely depend on overlapping retinotopic locations.

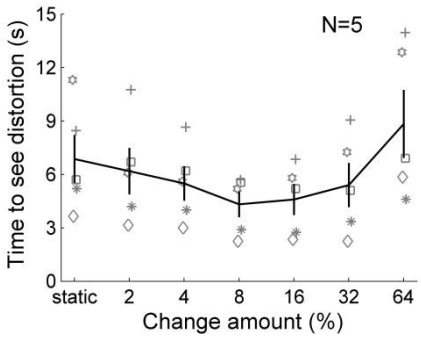

Figure 5.4 The time to see the distortion as a function of the amount of change across frames. The x-axis represents the base-2 logarithm of the amount of change except for the static condition (0% change). The solid line connects the means of five observers, and the error bars show $\pm 1\ SE$ across observers. Data for individual observers are shown with gray icons. Large but not excessively large variations between frames enhanced the distortion.

The distortion increase with a small amount of change suggests the role of signal uncertainty in the observed effect: with small variations among the external inputs the system is susceptible to the distortion. However, if the changes become excessive, the distortions decrease. One explanation is that large variations reduce the local effect by stimulating different retinotopic locations rather than constantly adapting the same location. Another possibility is that even for a global mechanism that may tolerate size changes but not necessary the aspect ratio difference, that is, the large change of aspect ratio alone reduced the distortion. To dissociate the effect of size and aspect ratio change, we generated a demo with the size change only (the same aspect ratio for all ellipses). The distortion appeared to be weaker, but further evaluation is required. Given the current results and observations, strong distortions tend to rely on the shape similarity among the sequence.

Though excessively large changes significantly increased the response time, the distortion still occurred quite rapidly. Also, when tested on the demo with the size change only, a fairly stable distortion would eventually emerge, which was similar to the condition with excessively large changes. We think that when the width and height were randomly drawn from uniform distributions while the center location was fixed, a concentric retinotopic pattern may be adapted through time, and in turn caused local adaptations and consequently a stable distortion (Roach, Webb, & McGraw, 2008). In addition, even among the sequence with large changes, similar shapes may still exist causing the distortion. This is different from a traditional paradigm in which only two distinct shapes are used for adaptor and test, respectively.

Experiment 4. Location jittering between frames

Since all possible shapes from the previous experiment still overlaid a relatively constrained retinotopic location, to further test whether the distortion was retinotopically limited, the effect of center location jittering between frames was measured in the experiment. Particularly, we presented ellipses at random locations to reduce retinotopic overlaps. If the effect reflects adaptation of shape processing mechanisms, then the distortion should also be induced with shapes that share similar geometries independent of the retinal locations.

Three conditions were tested in this experiment. The *aligned* condition was the standard condition in which ellipses were centered at 3 °eccentricity with 1.4 °mean minor radius (mean aspect ratio as 4:3), the amount of change between frames was 16% of the mean size, and the stimuli were altered at 4 Hz frequency. The *shift* condition was only different from the previous condition on the center locations: the ellipse center

would be randomly located within a $1.8° \times 2.4°$ square which was centered at 3 ° eccentricity. The *morph* condition was similar to the shift condition, except that each frame would be smoothly morphed from the previous one, so with the perception of continuous change. Five trials were measured for each condition, and a total of 15 trials were randomly interleaved for each observer. Data from seven observers were collected in this experiment.

Figure 5.5 shows the response time for each condition. Two out of seven observers could not perceive the distortion in the condition with location jittering (the *shift* condition). A paired t-test (assuming the differences to be approximately normally distributed) was conducted to assess effects between conditions. The response time in the *aligned* condition was significantly smaller than that in the *shift* condition ($t(6) = -4.67, p = 0.0034$). When comparing the response time in the *aligned* condition with that in the condition with both jittering and morphing, the former was still shorter ($t(6) = -2.56, p = 0.043$). No significant difference was found between the *shift* and *morph* condition ($t(6) = 1.87, p = 0.11$). The distortion observed in the *morph* condition may be induced by neighboring morphed frames (which are similar in terms of both location and shape), rather than across far-away frames. Overall, the distortions were hard to transfer across locations, which contradicted the global shape-selective mechanisms.

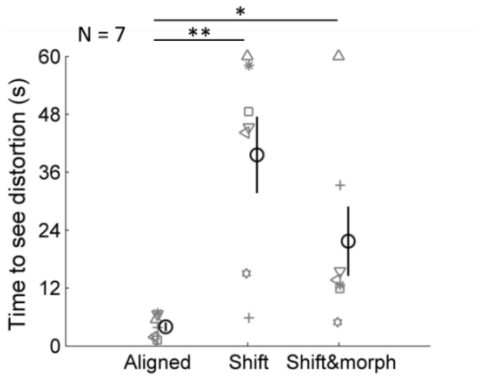

Figure 5.5 The response time measured in the three conditions with or without location jittering/morphing. The black circles show average of seven observers, and the error bars show $\pm 1\,SE$. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. The location shift significantly increases the response time, and limits the distortion.

Experiment 5. The role of eccentricity in shape distortion

In this experiment, we tested whether the absolute location of the presenting shape sequence would affect the distortion. Specifically, the response time was measured when the ellipse sequence was presented at different retinal eccentricities. If the effect is caused by global mechanisms, the response time should be similar among nearby locations, since neural responses to global shapes in higher level areas are shown to be location independent throughout the neurons' receptive field (M. Ito et al., 1995). We first tested conditions considering cortical magnification (Experiment 5.1), and in the second experiment (Experiment 5.2) the mean ellipse sizes were maintained across eccentricities to ensure that the eccentricity effect observed in Experiment 5.1 was not confound with the degree of curvature.

Experiment 5.1. Eccentricity

The stimulus eccentricity was manipulated in this experiment, and the stimulus size and line width were scaled for cortical magnification. Six conditions were tested: the center of stimulus was at 0.375 °, 0.75 °, 1.5 °, 3 °, 6 °, or 12 °eccentricity, and the mean length of corresponding minor radius was 0.192 °, 0.365 °, 0.71 °, 1.4 °, 2.78 °, or 5.54 °. The outline width was also changed accordingly. The amount of change between frames was 4% of the mean size of stimuli for each condition, and the stimuli were presented at 4 Hz alternating frequency. Five trials were measured for each condition, and a total of 30 trials were randomly interleaved for each observer. Data from seven observers were collected in this experiment.

Figure 5.6A shows the response time as we changed the eccentricities of scaled ellipses. It only took about 7 sec to see the distortion when large ellipses were presented at 12 degree eccentricity, but it took much longer when small ellipses were presented near the fovea. ANOVA was conducted to test the significance of the modulation of eccentricity. Since the data showed heterogeneity of variance (with Levene's test, $F(5,36) = 5.38, p < 0.001$), we applied the Welch correction to the denominator degrees of freedom in the *F*-test. This Welch ANOVA revealed that the response time was significantly different with various center eccentricities ($F(5,15.7) = 7.46, p < 0.001$).

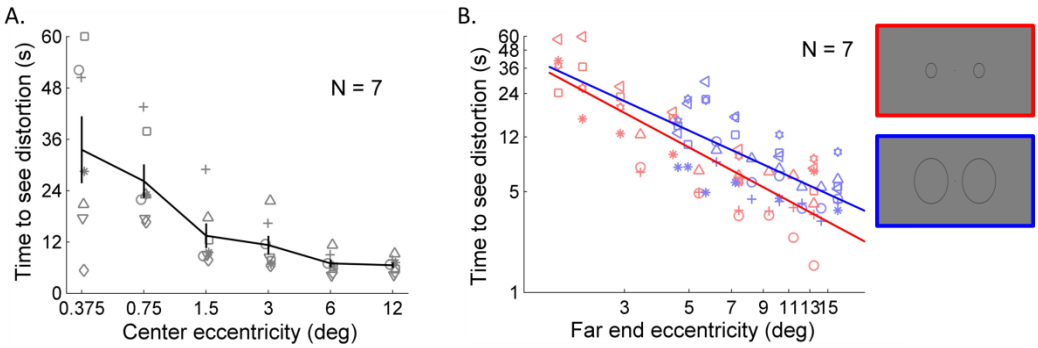

Figure 5.6 The response time as a function of the stimulus eccentricity. A. The response time when scaled ellipses were presented at different eccentricities. The x-axis represents the base-2 logarithm of the eccentricity of stimulus center. The error bars indicate $\pm 1\ SE$ of the response time across observers at each stimulus eccentricity. B. The response time when tested on ellipses with two different sizes at various eccentricities. Data are aligned to far end eccentricity for each condition. The red icons show the results from small testing ellipses, and the blue are from larger ellipses. The values of response time were log transformed to ensure homogeneity of error variances. Solid lines represent log-linear fit based on the measured data. No significant difference was observed between two lines.

Experiment 5.2. Degree of curvature

In the previous experiment we changed the size of ellipses while the eccentricity was manipulated, but in this experiment we fixed the size and wanted to look at whether the degree of curvature would have any influence on the effect of distortion. We tested the response time on ellipses with two different sizes at various eccentricities: the smaller ellipses (mean length of minor radius 1.4 °) had larger curvature than the larger ones (4.2 °). The ellipses were centered at various eccentricities. Since the eccentricity strongly influences the response time, for a fair comparison, we aligned the response time

according to the eccentricity of the stimulus far end (Figure 5.6B). For example, if the mean length of the ellipse minor radius was 1.4 °and the ellipse was centered at 3 ° eccentricity, the data would be plotted at 4.4 °along the x-axis. Analysis of covariance (ANCOVA) was used to test whether population means of the response time were equal across conditions with large/small ellipses, while statistically controlling for the effects of the far end eccentricities (as covariates).

Since the data again showed heterogeneity of variance (with Levene's test, $F(19,64) = 2.80, p = 0.001$), we first applied Box-Cox type transformations to stabilize variance. An optimized Box-Cox parameter ($\lambda = 0$) was acquired using boxcox (MASS) in R (2014), which indicates a logarithmic transformation of the response time would promote equality of the variance. With this transformation, the assumption, homogeneity of error variances, in ANCOVA was satisfied (with Levene's test, $F(19,64) = 0.53, p = 0.94$). According to the ANCOVA test, neither the slopes nor the intercepts of the small/lager curvature conditions were significantly different from the average estimates over all groups ($p = 0.40$ for slopes, and $p = 0.97$ for intercepts), which are shown in Figure 5.6B.

For another way to see the effect of curvature, we could directly see which part of the ellipse would start to show the distortion sooner – for example, whether the end near the minor axis (large curvature) or the one near the major axis (small curvature) tends to first show the distortion. Observers failed to report any particular order for the emergence of perceived polygons. The observation here was consistent with our results on no significant effect of degree of curvature. It has been shown that a strongly curved or a

weakly curved stimulus did not seem to give a significantly different curvature after-effect either (Gibson, 1933).

Together, the response decrease measured in Experiment 5.1 was mainly due to the increase of eccentricity rather than the decrease of curvature, and it cannot be explained by cortical magnification. A similar response decrease as the eccentricity increasing has been observed by Khuu et al. (2002) using the static circles. The decrease of receptor density and the increase of spacing with eccentricity (Curcio & Allen, 1990; Watson, 2014) may cause the ellipses susceptible to distortions. Interestingly, Khuu et al. (2002) also reported that when the same circle was presented, the number of illusory polygonal sides decreased as the eccentricity increasing. This result can be explained by receptive field size increase with the eccentricity (Jeremy Freeman & Simoncelli, 2011; Gattass, Gross, & Sandell, 1981; Gattass, Sousa, & Gross, 1988). To measure the correlation coefficient between the ellipse eccentricity and reported number of sides may help to locate the exact area responsive to the observed distortion.

Experiment 6. Spatial frequency

Our results until this point suggest that the distortion we observed may not reflect activity from mechanisms for global shape processing, and in order to evaluate the role of local processing mechanisms, we measured response time when ellipses were presented at various spatial frequency (Experiment 6) or contrast (Experiment 7). Lower level visual process (e.g., in the primary visual cortex) is known to be sensitive to contrast (Albrecht & Hamilton, 1982) and tuned to spatial frequency (De Valois, Albrecht, & Thorell, 1982; Ware & Mitchell, 1974). If the effect is independent of contrast or spatial frequency features, it would reflect activity from mechanisms beyond the lower level,

whereas dependency of these features should suggest mechanisms responsible for contrast or spatial frequency tuning at the low level. In this experiment, we first tested the effect of spatial frequency.

The spatial frequency of ellipse contour was controlled with a cross-sectional luminance profile defined by a radial fourth derivative of a Gaussian as suggested in Wilkinson, Wilson, and Habak (1998). Four spatial frequency conditions were tested: 1.5, 3.0, 6.0, and 11.3 cpd (Figure 5.7). Other parameters (mean ellipse size, eccentricity, change amount, and alternating frequency) were the same as in the standard condition, and the same contrast (56% Michelson contrast) was maintained across the four conditions. The response time to see the distortion was measured, and repeated in five trials for each condition.

A one-way ANOVA revealed a significant effect of spatial frequency on the measured response time ($F(3,119) = 7.14, p = 0.0033$), with the shortest response time after adapting to the highest spatial frequency pattern (Figure 5.7). Post hoc tests revealed that the response time in two high spatial frequency conditions (6.0 and 11.3 cpd) are significantly shorter than the other two conditions with peak frequency 1.5 and 3.0 cpd ($p < 0.05$ in each case, Tukey's HSD). Therefore, the spatial frequency affected the distortion: the polygonal perception occurred faster with higher spatial frequency contours than that with lower spatial frequency contours.

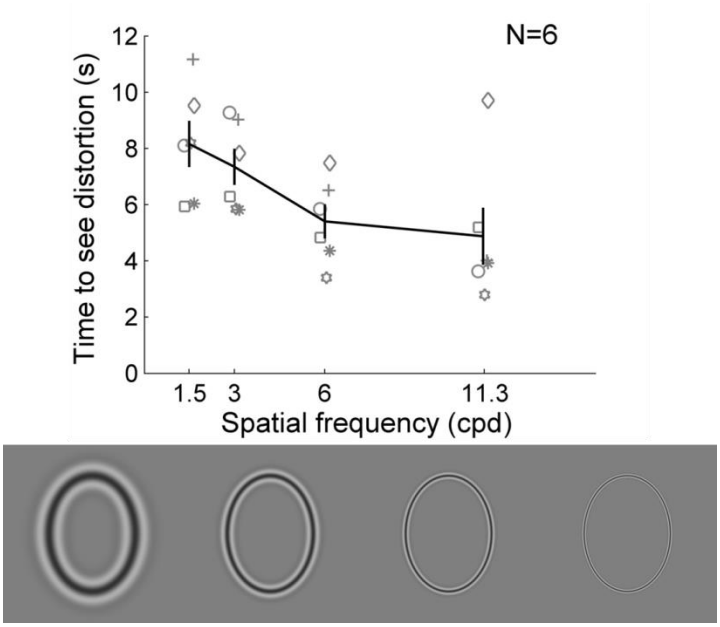

Figure 5.7 The example stimuli and results when the contour spatial frequency was manipulated. The y-axis indicates the peak spatial frequency. The x-axis represents the base-2 logarithm of the contour spatial frequency. From 1.5 cpd to 11.3 cpd (left to right), four conditions were tested.

Experiment 7. Contrast manipulation and interocular transfer

Similar as spatial frequency, responses in the striate cortex are highly modulated by stimulus contrast (Albrecht & Hamilton, 1982). Contrast normalization is largely thought to be accomplished in early cortical regions (Heeger, 1992), and responses from higher levels are reported to be less contrast dependent or rapidly saturated at low contrast energy (Avidan et al., 2002; Tootell et al., 1995). If the effect is mainly operated at higher levels than the striate cortex, one might expect to see the distortion insensitive to stimulus contrast. In the following experiments we further explored low-level feature dependency of the observed distortion. First, sensitivity to contrast modulation was examined (Experiment 7.1); additionally, the contribution of binocular processing to the

shape distortion was measured in the second part to rule out an exclusively monocular mechanism (Experiment 7.2).

Experiment 7.1. Binocular presentation

We tested whether the stimulus contrast at the adaptation phase would affect the strength of distortion for the test frames. Three levels of contrast, 12%, 54%, and 96% Weber contrast, were used for adapting ellipses. The contrast of the test frames were always at the intermediate level, and the adapting frames were at low-, intermediate-, or high-contrast level, therefore, three conditions were tested in total (Figure 5.8A). Between the 1st and 3rd condition, the relative contrast difference between the adaptor and test was the same, so we were able to compare the effect induced by low-contrast versus high-contrast adaptors while not confounded by the dissimilarity between the adaptor and test frames. We would focus on comparison between these two conditions.

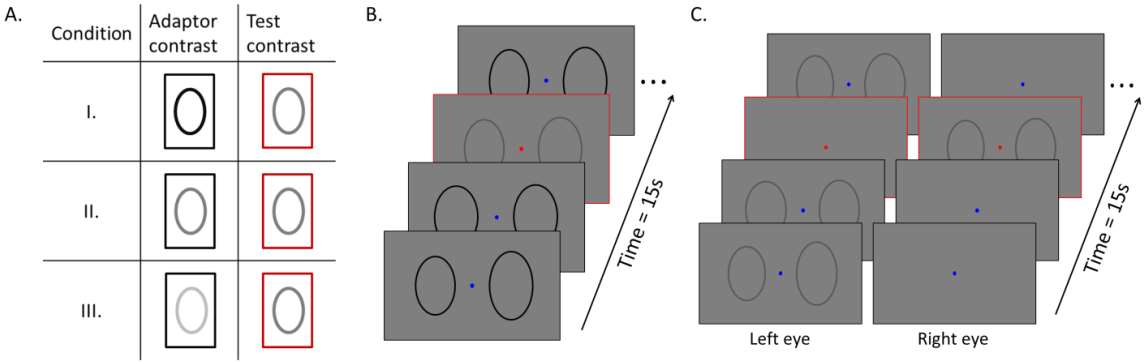

Figure 5.8 Experimental conditions and paradigms used to measure effects of contrast and Interocular transfer of the distortion. A. A summary of three conditions. B. Example of one trial, in which high-contrast ellipses were adapted, and intermediate-contrast ellipses were tested in Experiment 7.1. The red frame indicates the cued test frame. Each trial was 15 sec. A total of 15 trials were tested for each condition. C. Example

stimulus trial in Experiment 7.2. Adaptors and test frames were presented to different eyes.

The observers were instructed to pay attention to the entire sequence but only to report the perceived shape (distorted/polygon or not/ellipse) for a cued testing frame (the frame with a red fixation point instead of blue, see Figure 5.8B). One response option, *disappearing*, was also included in case the test frame was masked by the previous frame. In order to obtain information about temporal features of the emergence of the distortion, the test frames were randomly located in time. In detail, each trial was 15 sec, that is, 30 different ellipse frames with a 2Hz presentation (500 msec for each frame to ensure a doable task), and each frame would have a 20% probability to be a test frame, which would be cued using a red fixation point. Fifteen trials were measured for each condition, hence 45 trials in total for three conditions, and they were randomly interleaved. A 20 sec mask noise was inserted after each trial. We kept the same geometry for the ellipses as in the standard condition: at 3 °eccentricity with 1.4 °mean minor radius and with a 4% change amount across frames.

Figure 5.9A shows example data from one observer. Responses to the test frames were binned into 1.5 sec intervals, and the percentages of reported ellipse or polygon were calculated for each interval. The switch from perceiving ellipses to polygons was observed in all three conditions, and an earlier switch in condition I with high contrast adaptors was noticed. A small percentage of *disappearing* was also experienced by observers, especially in condition I with the higher contrast adaptor. To ensure fair comparison across conditions in following analysis, percentage of distortion was

calculated only among visible frames. Figure 5.9B shows summary results from seven observers. When averaged across the entire 15 sec, the condition with the same adaptor and testing contrast showed a significantly larger amount of reported distortion than the mean of the other two conditions with different adaptor/testing contrast ($t(12) = 3.80, p = 0.0025$); there was no difference between the conditions with high versus low contrast adaptors. However, this difference was significant during an early phase (Figure 5.9B lower panel, $t(12) = 3.44, p = 0.0049$) indicating that with higher contrast adaptors the induced distortion occurred sooner.

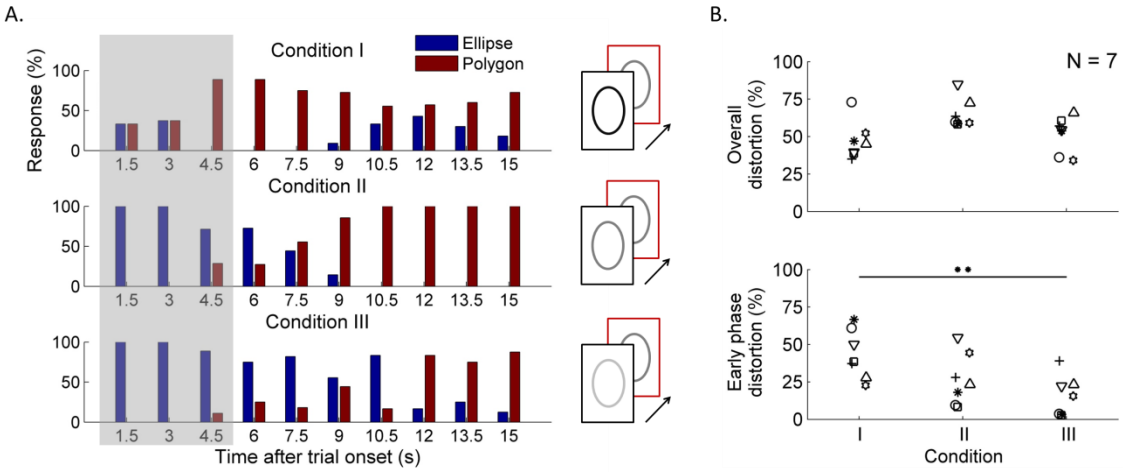

Figure 5.9 The results in Experiment 7.1 with binocular presentation. A. Example data from one observer. Shaded area indicates an arbitrary early phase that lasts 4.5 sec after the trial onset. B. Comparison between overall distortion and the distortion from the early phase indicated in A for seven observers. The y-axis indicates the percent polygonal distortion examined in three conditions with different adaptor contrast (from left to right, adapt to high, intermediate, and low contrast ellipses, respectively). Above: the distortion percentage for the entire trial (15 sec). Below: the distortion percentage for early phase of the trial (first 4.5 sec). During the early phase, a larger amount of

distortion was observed when adapting to high-contrast ellipses. Individual icons are for individual observers.

Experiment 7.2. Dichoptic presentation

To avoid masking among frames (especially in condition I with the higher contrast adaptor) and also to test interocular transfer of the effect, a dichoptic presentation was used in Experiment 7.2: adapting frames were presented to one eye, and testing frames to the other eye (Figure 5.8C). For each trial, the left or right eye was randomly selected for presenting the test frames. If the distortion relies on monocular mechanisms, presenting adaptor and test to different eyes would eliminate the effect; if the distortion remains, it would be most likely due to binocular mechanisms.

The summary results with the dichoptic presentation are presented in Figure 5.10A. A large amount of distortions were reported suggesting binocular transfer of the effect. The increase of overall distortion from similarity between the adaptor and testing contrast disappeared when the adaptor and test frames were presented to different eyes (Figure 5.10A upper panel, $t(12) = -0.98, p = 0.35$). During the first 4.5 sec, the difference between the conditions with high versus low contrast adaptors was again significant (Figure 5.10A lower panel, $t(12) = 3.44, p = 0.005$) replicating our observation with binocular presentations. The choice of the first 4.5 sec as the "early phase" so far was arbitrary, and in the next part of the analysis, we varied windows of this "early phase", which could be as short as 1.5 sec or as long as 15 sec. As the red curves in Figure 5.10B, across all tested early phases, the condition I appeared to have higher percentage of reported distortions than that in the condition III, and the t-test *p*-

values indicated a significant difference within the range of 3 to 10 sec. In a very brief time window, both conditions may show a small amount of distortions, so the difference was expected to be small; and when the window got too long, both conditions started to show stable distortions, the difference would decrease. The same analysis was performed with the binocular data in Experiment 7.1 as well (in blue).

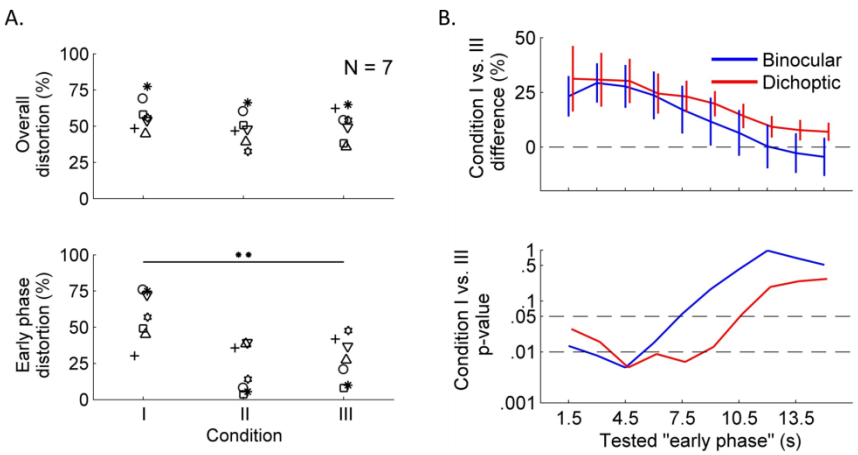

Figure 5.10 Results summary of Experiment 7. A. The percentages of reported distortion over the entire 15 sec and the distortion from the early phase (first 4.5 sec) with the dichoptic presentation. During the early phase, a significantly larger amount of distortion was reported in condition I than in condition III. B. The difference of reported distortion between the condition I and III as a function of tested “early phase” with varying length for both binocular (blue) and dichoptic (red) presentation. A positive number indicates more distortions reported in condition I than in condition III. The test *p*-values for the difference are shown in the lower panel.

The distortion with the high contrast adaptor showed a significantly larger amount of distorted perception with both binocular and dichoptic presentations at the early phase, which suggests an earlier distortion onset, presumably a stronger effect, when adaptors have a higher contrast. In a lower level adaptation effect, e.g., contrast/orientation

adaptation, high contrast adaptors do produce stronger effect than the one with low contrast (Georgeson & Harris, 1984; Harris & Calvert, 1989). Our results agree with the assumption that this distortion does not occur at high-level cortical areas whose responses to contrast tend to saturate fast (thus less sensitive to contrast), but rather at a lower level. Additionally, the effect is not exclusively monocular or simply due to retinal bleaching, since the interocular transfer of the distortion has been observed. This is consistent with Sakurai and Beaudot (2015), in which similar response time (the time to see the distortion) was acquired in the dichoptic and monocular conditions.

Experiment 8. Alternating frequency

Temporal features also play an essential role in adaptation. Two temporal parameters, alternating frequency and temporal gap, were manipulated in the following two experiments to explore the effect sensitivity to temporal features.

The same single geometry, ellipses at 3 °eccentricity with 1.4 °mean minor radius, was used, and with a 4% change amount across shape frames. Six conditions with alternating frequencies at 2 Hz, 4 Hz, 6 Hz, 10 Hz, 12 Hz, and 20 Hz were tested using the LCD monitor initially. Five repeat trials were measured for each condition, and a total of 30 trials were randomly interleaved for each observer. Data from six observers were collected. In a separate experiment, we tested the similar manipulation using the CRT monitor to ensure a more accurate timing. One static condition as in Experiment 3 was also included. Seven conditions (static, 1 Hz, 2 Hz, 5 Hz, 11 Hz, 21 Hz, and 42 Hz), hence 35 total trials, were tested for each observer, and five observers participated.

Figure 5.11A and 5.11B show the response time as a function of alternating frequency measured with the LCD monitor and CRT monitor, respectively. A one-way

ANOVA was conducted to test the significance of the main effect – alternating frequency: for the LCD results, $F(5,35) = 5.87, p = 0.0010$; for the CRT results, $F(5,29) = 5.37, p = 0.0027$. A longer response time appeared to be correlated with a lower alternating frequency. We also observed a trend that the observers spent longer time to see the distortion with the static presentation than that with the changing ellipses. The population marginal means of the static condition (16.55 sec) were significantly larger than the means of the conditions with various ellipses present ($p < 0.05$).

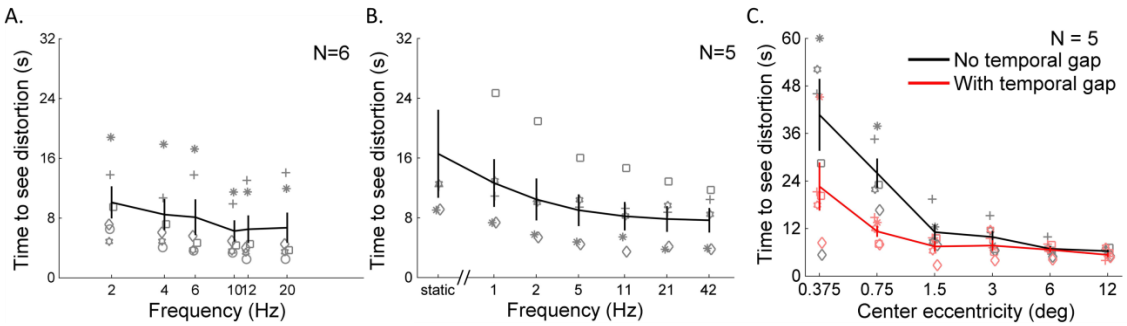

Figure 5.11 The response time in Experiment 8 and 9. The response time as a function of alternating frequency measured with the LCD monitor (A) and CRT monitor (B). The x-axis represents the base-2 logarithm of the amount of change except for the static condition. Both A and B indicate a response time decrease as the frequency increasing. C. The response time as a function of retinal eccentricity with (red) or without (black) temporal gaps between frames in Experiment 9. The marginal response time in the conditions with the temporal gap was significantly shorter than those in the no-gap conditions.

Experiment 9. Temporal gap between frames

In this experiment, we wanted to test whether the distortion would be reduced by inserting a temporal gap between frames. In addition to the conditions with the

eccentricity modulation described in Experiment 5.1, we included comparable conditions with a brief blank frame between successive frames. In detail, for the temporal gap conditions we maintained the alternating frequency at 4 Hz, but during each presentation period (250 msec), the stimulus was presented only for the first 117 msec followed by a blank frame for 133 msec (the fixation marker was remained on the screen the whole time).

With the temporal gap, a response time decrease was observed as shown in Figure 5.11C. Since the group means and variances were correlated, a logarithmic transform were applied to the response time (Levene's test after the transformation, $F(11,48) = 0.64, p = 0.78$). A two-way ANOVA (with/without temporal gap × six center eccentricity) was then performed. In addition to the center eccentricity ($F(5,59) = 15.40, p = 0$), a significant main effect was observed for the temporal gap ($F(1,59) = 19.51, p = 0.0038$). The marginal means of the response time in conditions with the temporal gap were significantly shorter than those in the no-gap conditions. No interaction between the eccentricity and temporal gap ($p = 0.42$) was shown. Consistently, an ANCOVA procedure gave significantly different fits for the intercepts ($p = 0.0013$), but not the slopes ($p = 0.12$).

The distortion persists with the temporal gap inserted, and the gaps lead to brief presentation durations that are actually similar as ones in the high frequency presentations. These results in Experiment 8 and 9 are expected according to the results that shorter test phase durations tend to induce stronger adaptation effect in a traditional aftereffect paradigm (Calvert & Harris, 1988; Wolfe, 1984). Note that though the

presentation time for each frame is brief, this is different from the brief adaptation that targets effects at a higher level (Miyashita & Chang, 1988; Suzuki, 2001). The adaptation time in our case should be the same as the induction time, which is about several seconds varying with experimental conditions.

5.3 Discussion

We have demonstrated that a robust polygonal shape distortion can be induced by adapting the visual system to a sequence of similar ellipses. The observed distortion is sensitive to low-level features such as contrast and spatial frequency, excessively large location or size changes across frames reduce the effect, and the distortion along the opposite direction (perceiving ellipses from polygons) almost does not occur. This distortion is different from the shape afterimage reported by H. Ito (2012). First, the perceived polygons are in the same contrast polarity as the presented ellipses, rather than in the opponent color as predicted by negative afterimages. Second, successful interocular transfer of the distortion indicates it is less likely due to the monocular mechanism or retinal bleaching, and it suggests a cortical locus for the perceptual effect. Additionally, even the static elliptical shape can induce this distortion and the observation that the distortion remains with exactly the same ellipse flashing indicate that the effect does not rely on apparent motion. Taken together, we think the effect mainly arises from local interaction likely in the primary visual cortex (V1). The global feature of a closed shape may strengthen the illusory polygonal perception. Here, we further discuss the roles of local and global mechanisms in this effect, and a perceptual model based on adaptation in V1 neurons is proposed. We then look at the role of input signal uncertainty

(mainly the influence of the shape variations, far eccentric and brief inputs). Finally, the possible functional roles and methodology benefits are discussed.

Local and global mechanisms

The distortion is initially caused by local adaptation, and the global factor could potentially enhance the closed polygonal shape perception, but it is not a primary cause. The perceived discontinuities (edges and corners) can be seen with flashed curves (e.g., a quarter of the ellipse contour), in which no closed shape exists. It is likely that the adaptation of oriented filters suppresses filtering responses to a local region (Marlin, Douglas, & Cynader, 1991; J. Anthony Movshon & Lennie, 1979), which causes discontinuities along the curve. Another possible local distortion is curvature adaptation (Gibson, 1933), in which prolonged viewing of, say, a convex curve would make the subsequent curve appear to be less convex. The decrease of curvature caused by local curvature adaptation can also induce discontinuities in visual perception. Since people are very sensitive to curvature discrimination (Watt & Andrews, 1982; Wilson, 1985, 1986) and coherence of circular contours (Levi & Klein, 2000; Wilkinson et al., 1998), these discontinuities are easy to be detected. The detected incomplete sampling or the change of curvature along an original smooth curve, together with a global pooling process, consequently causes the perception of a polygonal approximation of the ellipse. A schematic summary of the contributions of local and global process is shown in Figure 5.12.

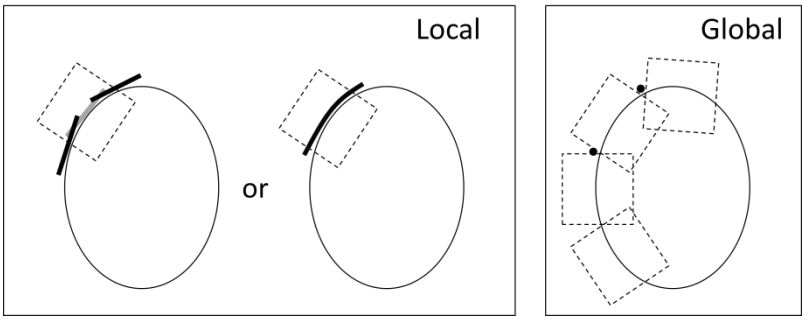

Figure 5.12 A schematic explanation for the observed distortions. The distortions arise from local orientation or curvature adaptation, and a global pooling process enhances the illusory polygonal perception. The dash squares indicate local regions.

Our results are not able to distinguish the initial cause as the oriented filter adaptation from the curvature adaptation. In fact, the existence of curvature detector has been a hot discussion (Blakemore & Over, 1974; MacKay & Mackay, 1974; Riggs, 1973; Sigel & Nachmias, 1975; Stromeyer & Riggs, 1974). The curvature information can be conveyed by the endstopped neurons (selective for orientation and stimulus length) found in cortical areas V1 and V2 (Dobbins, Zucker, & Cynader, 1987, 1989; Koenderink & Richards, 1988), or be encoded by multiple orientation-selective units displaced along the curve (Blakemore & Over, 1974; Poirier & Wilson, 2006; Sigel & Nachmias, 1975; Wilson & Richards, 1989). However, the existence of specialized curvature detector has also been demonstrated by dissociate the local orientation effect from remaining effects that are potentially associated with a higher order processing unit (Bell, Gheorghiu, Hess, & Kingdom, 2011; Gheorghiu & Kingdom, 2006, 2007a, 2007b, 2008; Hancock & Peirce, 2008; Keeble & Hess, 1999). Our results can be explained by adapting line-orientations, and may also be explained using adaptation with the curvature processing units. Further work is needed to isolate these two local mechanisms.

It is evident that with the observed distortion local adaptation is dominant, but our results may not firmly rule out the effect of global mechanisms. In fact, the remaining effects in the *shift* condition in Experiment 4, the observation with the size varied ellipse sequence inducing similar distortions (after Experiment 3), and the intermittent perception of ellipses with polygon presentation from one out of six observers in Experiment 2 may suggest a weak adaptation at a shape-selective level. The interaction between the local and global process could be fairly complex. A local effect can be inherited downstream by higher levels (Dickinson, Almeida, Bell, & Badcock, 2010; Dickinson, Harman, Tan, Almeida, & Badcock, 2012; Kohn & Movshon, 2003); extra high-level adaptation can also occur on top of the low-level effect (N. D. Anderson et al., 2007; Xu, Dayan, Lipkin, & Qian, 2008; Xu, Liu, Dayan, & Qian, 2012); adapting to global pattern may induce local aftereffect as well (Roach et al., 2008); additionally, the relative importance of low-level and high-level effects may not be fixed, for example, perceptual grouping will increases the global adaptation while decreasing the local adaptation (He, Kersten, & Fang, 2012). Understanding the roles of the local/global process would help understand functions of visual processing along the hierarchy.

V1 model simulation

To get a concrete idea about the effect of local adaptation, especially the oriented filter adaptation in the observed distortion, we simulated the incomplete sampling as a consequence of the local adaptation using a V1 response model. The V1 filter bank defined in frequency domain was composed of sixteen evenly distributed orientation channels and six spatial frequency channels peak at 0.75, 1.5, 2.8, 4.4, 8.0 and 16.0 cpd

(Wilson, McFarlane, & Phillips, 1983) at phase 0 and $\frac{\pi}{2}$. A filter K sensitive to the orientation θ_i and spatial frequency SF_j at phase ϕ is labeled as $K(\theta_i, SF_j, \phi)$, where $i \in 1, 2, \dots, 16$, $j \in 1, 2, \dots, 6$, and $\phi = 0°, 90°$. Given an image input $s(x, y)$, the output of filter K can be written as $O\big(K(\theta_i, SF_j, \phi), s(x, y)\big) = K(\theta_i, SF_j, \phi) \cdot \mathcal{F}(s(x, y))$, where $\mathcal{F}(s(x, y))$ is the Fourier transform of the input.

The adaptation effect was simulated by suppressing filtering responses based on response energy E_{SF_j} for each spatial frequency channel, where $E_{SF_j} = \sum_\theta O\big(K(\theta_i, SF_j, \phi), s(x, y)\big)$. In detail, if the response energy is large in certain spatial frequency channel, the suppression would be strong. The suppression coefficient for each spatial frequency channel is defined as $\rho_j = \frac{E_{SF_j}}{\sum_{SF} E_{SF_j}}$. Within each spatial frequency channel, the number of suppressed orientation channels was determined by the suppression coefficient (round $16 \times \rho_j$ to the nearest integer), and then the suppressed channels were randomly selected. The energy at the orientation channels was quite similar given the circular input, and random selection simulated a stochastic process. Perceptually, these random "decisions" may be enhanced by global interpretations that vertices and polygon appear. The suppressed output O' was transformed back to the spatial domain, and the complex cell response was calculated based on the quadrature phase simple cell responses, $\sqrt{O'(\phi = 0°)^2 + O'(\phi = 90°)^2}$. The simulated final output as the sum of complex cell responses at various orientations and spatial frequencies is shown in Figure 5.13B or C. The model was further applied to stimuli used in Experiment 6 when the spatial frequency of contours were manipulated as shown in

Figure 5.13 D-I. The simulated distortions predict a weaker distortion with the low spatial frequency contour and a stronger distortion with the high spatial frequency contour. This is also consistent with the results presented in Figure 5.7. The receptive field sensitive to lower spatial frequency tends to cover a larger spatial space than the one for higher spatial frequency. This larger receptive field gives tolerance to incomplete sampling caused by adaptation. Therefore, given the same receptive field density and the same amount of adaptation, the distorted output representation would be harder to be noticed for the lower spatial frequency pattern. As for the ellipses presented at larger eccentricity, larger receptive fields are also expected there, but they tend to be at a lower density as well, which may eventually leads to stronger distortions at far eccentricity. This is also consistent with the predictions based on perceptual uncertainty, which will be discussed next.

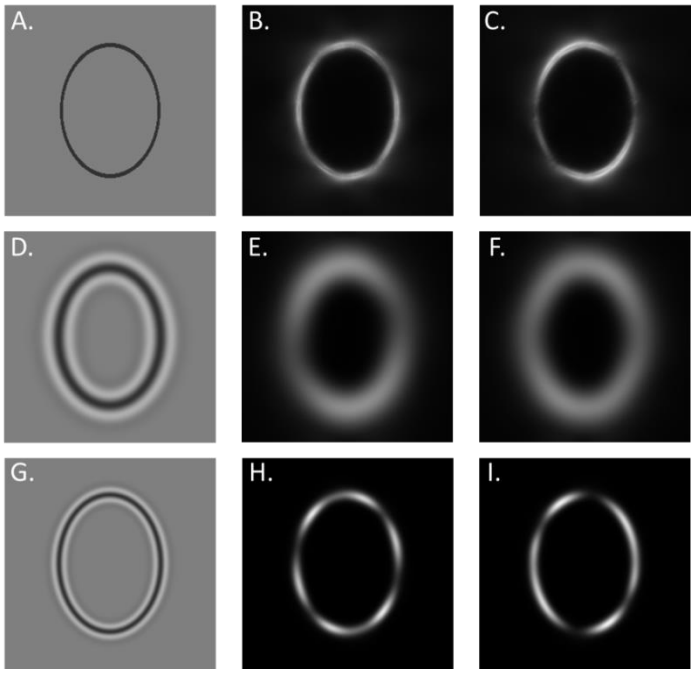

Figure 5.13 Examples of simulated distortion based on V1 adaptation model. A-C. Simulation of one ellipse (A) from the standard condition. B and C are two simulated examples. They are slightly different due to the random selection of suppressed channels. D-F. Simulation of an ellipse with low spatial frequency contour (1.5 cpd). G-I. Simulation of an ellipse at a higher spatial frequency (6 cpd).

The role of perceptual uncertainty

In Experiment 3, the conditions with small shape variations induced a stronger distortion than the static condition (the condition without variation); the distortion in the conditions with large eccentricities was shown to be stronger than those in the conditions closer to the fovea in Experiment 5; additionally, an increase was observed with brief presentations for individual frames in Experiment 8 and 9. These results together suggest that perceptual uncertainty to the input would enhance the illusory perception. We will elaborate this idea from three parts. First, this uncertainty is not necessary to be predicted

by the amount of signal, though they could be related. For example, in the case with shape variations, the amount of signal can be quite similar for large variation sequence and small variation sequence, but uncertainty in the sequence with large variations is substantial comparing with the low variation one. However, in the case with brief presentations, the amount of signal largely determines certainty about the shape. As for the eccentricity, though it is possible to match peripheral vision performance in various tasks by enlarging peripheral stimulus to potentially match signal amount to the system, the elevation of distortions in peripheral after considering cortical magnification has been observed in many cases as well. For example, the contextual distortions with lower-level features, such as contrast (Snowden & Hammett, 1998; Xing & Heeger, 2000) and orientation (Mareschal et al., 2010; Muir & Over, 1970; Over, Broerse, & Crassini, 1972; Solomon et al., 2004) increase in peripheral; and higher-level percepts, such as shape (Suzuki & Cavanagh, 1998) and letter recognition (Bouma, 1970) are more severely influenced. It is possible that the perceptual uncertainty in peripheral cannot be fully accounted by enhancing the input.

Second, many previous studies have emphasized a similar role of uncertainty or visibility in influencing perceptual judgement (Calvert & Harris, 1988; Gardelle, Kouider, & Sackur, 2010; Mruczek, Blair, Strother, & Caplovitz, 2015; Suzuki & Cavanagh, 1998). In an orientation matching task, Gardelle et al. (2010) simultaneously acquired data of matched orientation and reported level of visibility, and they found that the orientation bias was the largest when the visibility was at the intermediate level, which suggests that subjective uncertainty (up to a certain point) increases perceived

orientation deviation. This is exactly what we observed in experiment 3. Functionally, this stronger distortion associated with larger perceptual uncertainty may further serve the functional goal of better representing complex information, especially when the distortion is taken as compensation for otherwise easily confused or degraded inputs (Qiu, Kersten, & Olman, 2013).

Last, larger perceptual uncertainty likely increases the level of system internal noise which excites more responsive detectors for various features, and over time causes adaptation involving a larger number of feature detectors. Adaptation is a continuous process, and further adaptation in partially suppressed channels would be harder and slower than in channels that have not been adapted (Harris & Calvert, 1989; Magnussen & Greenlee, 1985). Therefore, adapting to a broad band distribution seems to be more effective by avoiding adaptation saturation (when the overall amount of adaption is reasonably large). This may be confused with results in which adaptation to a broad band distribution often induces a weaker effect comparing with a very narrow band adaptation (Dekel & Sagi, 2015; Goddard et al., 2008; McGovern, Roach, & Webb, 2014; Price & Prescott, 2012). The major difference between these results and the results we present here is that the former usually adapting to one value but testing on a different value, but here no distinct boundary between adaptor and test stimulus.

Functional roles and methodology significance

The visual system functionally creates abstraction to optimize representations (Attneave, 1954), and this could be particularly important when multiple objects need to be encoded at a short period of time. The visual system continually adapts to the environment dynamics potentially to improve its function (Barlow, 1991; Clifford et al.,

2007; Kohn, 2007; Wark, Lundstrom, & Fairhall, 2007; Webster, 2011). The visual misperception here may reflect an efficient representation maybe as a statistical template of rapidly changing external inputs, and we showed that this process happens very early in the visual processing stream.

Adaptation also has been used intensively as a tool to study mechanisms that support visual processes. Selective adaptation has been observed in various low-level features such as contrast (Blakemore & Campbell, 1969; Kohn, 2007), orientation (Clifford et al., 2007; Gibson & Radner, 1937), and spatial frequency (Blakemore, Nachmias, et al., 1970), and these results have advanced our understanding of low-level visual processing mechanisms. Perceptual aftereffects with complex patterns (Alter & Schwartz, 1988; N. D. Anderson et al., 2007; Bell & Kingdom, 2009; H. Ito, 2012; Regan & Hamstra, 1992; Suzuki, 2001, 2003; Suzuki & Cavanagh, 1998; Webster & Maclin, 1999) suggest that higher levels of the visual system, such as a global shape-selective process, can also adapt (Clifford et al., 2007; Webster, 2011). Such results further elucidate the response characteristics at higher levels of the visual hierarchy. Here, our study suggests a paradigm with fast repeated presentations to examine the effect of adaptation, which better simulates a dynamic environment and promotes a continuous visual process to effectively respond to this environmental regularity.

5.4 Conclusion

We observed polygonal perception with fast repeated presentation of ellipses. This effect is strongly influenced by local features, and manipulations maintaining global similarity are not enough to maintain the same amount of distortion. We think plasticity

occurring early in the processing stream, and then impact downstream perception. Further, perceptual uncertainty, for example due to certain amount of variations, large eccentricity, and short presentation duration, enhances the effect. This distorted perception may serve the functional goal of better representing complex information.

6. Summary

6.1 Conclusions

In this thesis I have described a set of experiments focusing on manipulating visual context from multiple levels, which demonstrate that the cortical responses of early visual areas are highly sensitive to various contexts. In particular, the global shape complexity, figure-ground segregation or perceptual group, and the uncertainty of target perception would influence perceptual biases and/or contextual modulations in early visual cortex. Such process potentially benefits visual functions by exaggerating local features especially when they would otherwise be compulsorily pooled with surrounding stimuli. We also modified a computational framework, originally developed by Schwartz et al. (2009), to quantitatively characterize effects from both local and long-range context. This framework can also be used to explore the source of variations in the contextual modulation, and its application is briefly discussed next.

6.2 Future directions

6.2.1 Local contextual modulation influenced by disease

It has been shown that diverse cognitive deficits in schizophrenia are due to an impaired ability to analyze contextual information (Silverstein & Schenkel, 1997), and such deficits have been observed in early visual processing as well (Butler, Silverstein, & Dakin, 2008). In particular, weaker surround suppression during contrast perception (Dakin, Carlin, & Hemsley, 2005; Schallmo, Sponheim, & Olman, 2015; Tadin et al.,

2006; Yoon et al., 2009) and contour integration deficit (Dakin & Baruch, 2009; Robol, Casco, & Dakin, 2012; Schallmo, Sponheim, & Olman, 2013) have been demonstrated in patients with schizophrenia, and the relationship among various contextual effects has also been discussed (Robol et al., 2013; Tibber et al., 2013; Yang et al., 2013). It is worth noting that the patients do not consistently show deficits among all contextual tasks, for example, they seem to show normal or even stronger contextual modulations with luminance, size, orientation or motion perception (Y. Chen, Norton, & Ongur, 2008; Tibber et al., 2013; Yang et al., 2013). It is likely that only certain levels of the contextual modulation are impaired but not the others. Our existing computational framework could potentially help to isolate the effect from various levels. Fully exploring the source of contextual impairments may provide us a better understanding of neuropathology in schizophrenia, or other mental disorders, such as bipolar and autism spectral disorder.

6.2.2 Feedforward and feedback interactions in the human visual cortex with depth-dependent fMRI

We have shown contextual modulations reflected in V1 responses from various local/global cues, and an extended goal is to distinctively identify these modulations with different characteristics at different cortical layers in V1. The cortex has a laminar structure with six layers spanning a depth of 2-4 mm. Forward projections arise mainly from superficial (supra-granular) layers and terminate mainly in middle (granular) layers; feedback projections arise from both superficial and deep (infra-granular) layers and terminate mainly outside middle layers (Felleman & Van Essen, 1991; Maunsell & Newsome, 1987; Rockland & Pandya, 1979; Zeki & Shipp, 1988). The rich set of

feedback projections may serve important functional roles in top-down modulations, but direct evidence is limited.

Experiments with conditions differentiating various levels of contextual modulations originated from long-range horizontal connections or inter-areal feedback effects (Angelucci & Bressloff, 2006; Angelucci et al., 2002; Bair, Cavanaugh, & Movshon, 2003; Webb, 2005) might show different depth-dependent response profiles in V1 (Shushruth et al., 2013). Using ultrahigh field fMRI with submillimeter resolution, we can record signals from different cortical depth. It has been hypothesized that middle layers would show local computations, whereas superficial and deep layers are susceptible to inter-areal feedback modulations (Friston, 2005, 2008). Further, communications or cross talks among areas are essential in cortical processes, and laminar organizations in cortex are suggested to be infrastructure for implementing such communication loop. Correlation analysis with depth-dependent fMRI would be useful methods for exploring communications happening among cortical regions.

Bibliography

Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, 300:523-525.

Albrecht, D. G., Farrar, S. B., & Hamilton, D. B. (1984). Spatial contrast adaptation characteristics of neurones recorded in the cat's visual cortex. *The Journal of Physiology, 347*(1), 713-739.

Albrecht, D. G., & Geisler, W. S. (1991). Motion selectivity and the contrast-response function of simple cells in the visual cortex. *Visual Neuroscience, 7*(06), 531-546. doi: 10.1017/S0952523800010336

Albrecht, D. G., & Hamilton, D. B. (1982). Striate cortex of monkey and cat: contrast response function. *Journal of Neurophysiology, 48*(1), 217-237.

Albright, T. D., & Desimone, R. (1987). Local precision of visuotopic organization in the middle temporal area (MT) of the macaque. *Experimental Brain Research*, 65(63), 582–592. doi:510.1007/BF00235981.

Allman, J., Miezin, F., & McGuinness, E. (1985). Stimulus Specific Responses from Beyond the Classical Receptive Field: Neurophysiological Mechanisms for Local-Global Comparisons in Visual Neurons. *Annual Review of Neuroscience, 8*(1), 407-430. doi: 10.1146/annurev.ne.08.030185.002203

Alter, I., & Schwartz, E. L. (1988). Psychophysical Studies of Shape with Fourier Descriptor Stimuli. *Perception, 17*(2), 191-202. doi: 10.1068/p170191

Altmann, C. F., Bülthoff, H. H., & Kourtzi, Z. (2003). Perceptual Organization of Local Elements into Global Shapes in the Human Visual Cortex. *Current Biology, 13*(4), 342-349. doi: 10.1016/S0960-9822(03)00052-6

Andersen, R. A., & Bradley, D. C. (1998). Perception of three-dimensional structure from motion. *Trends in cognitive sciences*, 2(6), 222–228.

Anderson, J. C., & Martin, K. A. C. (2009). The Synaptic Connections between Cortical Areas V1 and V2 in Macaque Monkey. *The Journal of Neuroscience, 29*(36), 11283-11293. doi: 10.1523/JNEUROSCI.5757-08.2009

Anderson, N. D., Habak, C., Wilkinson, F., & Wilson, H. R. (2007). Evaluating shape after-effects with radial frequency patterns. *Vision Research, 47*(3), 298-308. doi: 10.1016/j.visres.2006.02.013

Angelucci, A., & Bressloff, P. C. (2006). Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons *Progress in brain research* (Vol. 154, pp. 93-120): Elsevier.

Angelucci, A., Levitt, J. B., Walton, E. J. S., Hupé, J.-M., Bullier, J., & Lund, J. S. (2002). Circuits for Local and Global Signal Integration in Primary Visual Cortex. *The Journal of Neuroscience, 22*(19), 8633-8646.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological review, 61*(3), 183-193.
Avidan, G., Harel, M., Hendler, T., Ben-Bashat, D., Zohary, E., & Malach, R. (2002). Contrast Sensitivity in Human Visual Areas and Its Relationship to Object Recognition. *Journal of Neurophysiology, 87*(6), 3102-3116.
Bair, W., Cavanaugh, J. R., & Movshon, J. A. (2003). Time course and time-distance relationships for surround suppression in macaque V1 neurons. *The Journal of Neuroscience, 23*(20), 7690-7701.
Bandettini, P. A., Jesmanowicz, A., Wong, E. C., & Hyde, J. S. (1993). Processing strategies for time-course data sets in functional mri of the human brain. *Magnetic Resonance in Medicine, 30*(2), 161-173. doi: 10.1002/mrm.1910300204
Barlow, H. B. (1959). Sensory mechanisms, the reduction of redundancy, and intelligence. *NPL Symposium on the Mechanization of Thought Process, 10*, 535-539.
Barlow, H. B. (1961). Possible principles underlying the trasformations of sensory messages (pp. 217-234). Cambridge, MA: MIT Press.
Barlow, H. B. (1991). A theory about the functional role and synaptic mechanism of visual after-effects *Vision*: Cambridge University Press.
Barone, P., Batardiere, A., Knoblauch, K., & Kennedy, H. (2000). Laminar Distribution of Neurons in Extrastriate Areas Projecting to Visual Areas V1 and V4 Correlates with the Hierarchical Rank and Indicates the Operation of a Distance Rule. *The Journal of Neuroscience, 20*(9), 3263-3281.
Bauer, R., & Heinze, S. (2002). Contour integration in striate cortex. Classic cell responses or cooperative selection? *Experimental Brain Research, 147*(2), 145-152. doi: 10.1007/s00221-002-1178-6
Bednar, J. A., & Miikkulainen, R. (2000). Tilt aftereffects in a self-organizing model of the primary visual cortex. *Neural Computation, 12*(7), 1721-1740.
Bell, J., Gheorghiu, E., Hess, R. F., & Kingdom, F. A. A. (2011). Global shape processing involves a hierarchy of integration stages. *Vision Research, 51*(15), 1760-1766. doi: 10.1016/j.visres.2011.06.003
Bell, J., & Kingdom, F. A. A. (2009). Global contour shapes are coded differently from their local components. *Vision Research, 49*(13), 1702-1710. doi: 10.1016/j.visres.2009.04.012
Blakemore, C., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *The Journal of Physiology, 203*(1), 237-260.231.
Blakemore, C., Carpenter, R. H. S., & Georgeson, M. A. (1970). Lateral inhibition between orientation detectors in the human visual system. *Nature, 228*(5266), 37-39. doi: 10.1038/228037a0
Blakemore, C., Carpenter, R. H. S., & Georgeson, M. A. (1971). Lateral thinking about lateral inhibition. *Nature, 234*(5329), 418-419. doi: 10.1038/234418b0
Blakemore, C., Nachmias, J., & Sutton, P. (1970). The perceived spatial frequency shift: evidence for frequency-selective neurones in the human brain. *The Journal of Physiology, 210*(3), 727-750.

Blakemore, C., & Over, R. (1974). Curvature detectors in human vision? *Perception, 3*(1), 3-7. doi: 10.1068/p030003
Blakemore, C., & Tobin, E. A. (1972). Lateral inhibition between orientation detectors in the cat's visual cortex. *Experimental Brain Research, 15*(4), 439-440. doi: 10.1007/BF00234129
Born, R. T., & Bradley, D. C. (2005). Structure and Function of Visual Area MT. *Annual Review of Neuroscience*, 28(21), 157–189. doi:110.1146/annurev.neuro.1126.041002.131052.
Bosking, W. H., Zhang, Y., Schofield, B., & Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *The Journal of Neuroscience, 17*(6), 2112-2127.
Bouma, H. (1970). Interaction Effects in Parafoveal Letter Recognition. *Nature, 226*(5241), 177-178. doi: 10.1038/226177a0
Boyaci, H., Fang, F., Murray, S. O., & Kersten, D. (2010). Perceptual Grouping-Dependent Lightness Processing in Human Early Visual Cortex. *Journal of Vision, 10*(9). doi: 10.1167/10.9.4
Boynton, G. M., & Hegdé, J. (2004). Visual Cortex: The Continuing Puzzle of Area V2. *Current Biology, 14*(13), R523-R524. doi: 10.1016/j.cub.2004.06.044
Bradley, D. C., Chang, G. C., & Andersen, R. A. (1998). Encoding of three-dimensional structure-from-motion by primate area MT neurons. *Nature*, 392(6677), 6714–6717. doi:6610.1038/33688.
Brainard, D. (1997). The Psychophysics Toolbox. *Spatial vision, 10*(4), 433.
Brincat, S. L., & Connor, C. E. (2004). Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nature Neuroscience, 7*(8), 880-886. doi: 10.1038/nn1278
Bullier, J., Hupé, J. M., James, A., & Girard, P. (1996). Functional interactions between areas V1 and V2 in the monkey. *Journal of Physiology-Paris, 90*(3), 217-220.
Bushnell, B. N., Harding, P. J., Kosai, Y., & Pasupathy, A. (2011). Partial Occlusion Modulates Contour-Based Shape Encoding in Primate Area V4. *Journal of Neuroscience, 31*(11), 4012-4024. doi: 10.1523/JNEUROSCI.4766-10.2011
Butler, P. D., Silverstein, S. M., & Dakin, S. C. (2008). Visual Perception and Its Impairment in Schizophrenia. *Biological Psychiatry, 64*(1), 40-47. doi: 10.1016/j.biopsych.2008.03.023
Buzsáki, G., Kaila, K., & Raichle, M. (2007). Inhibition and brain work. *Neuron, 56*(5), 771-783. doi: 10.1016/j.neuron.2007.11.008
Calvert, J. E., & Harris, J. P. (1988). Spatial frequency and duration effects on the tilt illusion and orientation acuity. *Vision Research, 28*(9), 1051-1059. doi: 10.1016/0042-6989(88)90082-X
Carandini, M., & Heeger, D. J. (1994). Summation and division by neurons in primate visual cortex. *Science (New York, N.Y.), 264*(5163), 1333-1336.
Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience, 13*(1), 51-62. doi: 10.1038/nrn3136

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and Normalization in Simple Cells of the Macaque Primary Visual Cortex. *The Journal of Neuroscience, 17*(21), 8621-8644.

Cardin, V., Friston, K. J., & Zeki, S. (2010). Top-down Modulations in the Visual Form Pathway Revealed with Dynamic Causal Modeling. *Cerebral Cortex, 21*(3), 550-562. doi: 10.1093/cercor/bhq122

Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002a). Nature and interaction of signals from the receptive field center and surround in macaque v1 neurons. *J Neurophysiol*, 2530-2546.

Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002b). Selectivity and Spatial Distribution of Signals From the Receptive Field Surround in Macaque V1 Neurons. *Journal of Neurophysiology, 88*(5), 2547-2556. doi: 10.1152/jn.00693.2001

Chen, C.-C., & Tyler, C. W. (2001). Lateral sensitivity modulation explains the flanker effect in contrast discrimination. *Proceedings of the Royal Society of London B: Biological Sciences, 268*(1466), 509-516. doi: 10.1098/rspb.2000.1387

Chen, C.-C., & Tyler, C. W. (2002). Lateral modulation of contrast discrimination: flanker orientation effects. *Journal of Vision, 2*(6), 520-530. doi: 10.1167/2.6.8

Chen, C.-C., & Tyler, C. W. (2008). Excitatory and inhibitory interaction fields of flankers revealed by contrast-masking functions. *Journal of Vision, 8*(4). doi: 10.1167/8.4.10

Chen, M., Yan, Y., Gong, X., Gilbert, C. D., Liang, H., & Li, W. (2014). Incremental Integration of Global Contours through Interplay between Visual Cortical Areas. *Neuron, 82*(3), 682-694. doi: 10.1016/j.neuron.2014.03.023

Chen, Y., Norton, D., & Ongur, D. (2008). Altered Center-Surround Motion Inhibition in Schizophrenia. *Biological Psychiatry, 64*(1), 74-77. doi: 10.1016/j.biopsych.2007.11.017

Cheng, K., Fujita, H., Kanno, I., Miura, S., & Tanaka, K. (1995). Human cortical regions activated by wide-field visual motion: an H2(15)O PET study. *Journal of Neurophysiology, 74*(1), 413-427.

Clifford, C. W. G., & Harris, J. A. (2005). Contextual Modulation outside of Awareness. *Current Biology, 15*(6), 574-578. doi: 10.1016/j.cub.2005.01.055

Clifford, C. W. G., Spehar, B., Solomon, S. G., Martin, P. R., & Zaidi, Q. (2003). Interactions between color and luminance in the perception of orientation. *Journal of Vision, 3*(2). doi: 10.1167/3.2.1

Clifford, C. W. G., Webster, M. A., Stanley, G. B., Stocker, A. A., Kohn, A., Sharpee, T. O., & Schwartz, O. (2007). Visual adaptation: Neural, psychological and computational aspects. *Vision Research, 47*(25), 3125-3131. doi: 10.1016/j.visres.2007.08.023

Clifford, C. W. G., Wenderoth, P., & Spehar, B. (2000). A functional angle on some after-effects in cortical vision. *Proc Biol Sci, 267*(1454), 1705-1710. doi: 10.1098/rspb.2000.1198

Coen-cagli, R., Dayan, P., & Schwartz, O. (2009). *Statistical models of linear and nonlinear contextual interactions in early visual processing*.

Coen-Cagli, R., Dayan, P., & Schwartz, O. (2012). Cortical Surround Interactions and Perceptual Salience via Natural Scene Statistics. *PLoS Comput Biol, 8*(3). doi: 10.1371/journal.pcbi.1002405

Cornsweet, T. N. (1962). The Staircase-Method in Psychophysics. *The American Journal of Psychology, 75*(3), 485-491. doi: 10.2307/1419876

Cox, R. W. (1996). AFNI: Software for Analysis and Visualization of Functional Magnetic Resonance Neuroimages. *Computers and Biomedical Research, 29*(3), 162-173. doi: http://dx.doi.org/10.1006/cbmr.1996.0014

Curcio, C. A., & Allen, K. A. (1990). Topography of ganglion cells in human retina. *The Journal of Comparative Neurology, 300*(1), 5-25. doi: 10.1002/cne.903000103

Dakin, S. C. (2009). Vision: Thinking Globally, Acting Locally. *Current Biology, 19*(18), R851-R854. doi: 10.1016/j.cub.2009.08.021

Dakin, S. C., & Baruch, N. J. (2009). Context influences contour integration. *Journal of Vision, 9*(2), 13-13. doi: 10.1167/9.2.13

Dakin, S. C., Carlin, P., & Hemsley, D. (2005). Weak suppression of visual context in chronic schizophrenia. *Current biology: CB, 15*(20), R822-824. doi: 10.1016/j.cub.2005.10.015

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *NeuroImage, 9*(2), 179-194. doi: 10.1006/nimg.1998.0395

Das, A., & Gilbert, C. D. (1999). Topography of contextual modulations mediated by short-range interactions in primary visual cortex. *Nature, 399*(6737), 655-661. doi: 10.1038/21371

de-Wit, L. H., Kubilius, J., Wagemans, J., & Op de Beeck, H. P. (2012). Bistable Gestalts reduce activity in the whole of V1, not just the retinotopically predicted parts. *Journal of Vision, 12*(11). doi: 10.1167/12.11.12

De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research, 22*(5), 545-559. doi: 10.1016/0042-6989(82)90113-4

Dekel, R., & Sagi, D. (2015). Tilt aftereffect due to adaptation to natural stimuli. *Vision Research, 117*, 91-99. doi: 10.1016/j.visres.2015.10.014

DeYoe, E. A., Carman, G. J., Bandettini, P., Glickman, S., Wieser, J., Cox, R., . . . Neitz, J. (1996). Mapping striate and extrastriate visual areas in human cerebral cortex. *Proceedings of the National Academy of Sciences, 93*(6), 2382-2386.

DeYoe, E. A., Felleman, D. J., Van Essen, D. C., & McClendon, E. (1994). Multiple processing streams in occipitotemporal visual cortex. *Nature, 371*(6493), 151-154. doi: 10.1038/371151a0

Dickinson, J. E., Almeida, R. A., Bell, J., & Badcock, D. R. (2010). Global shape aftereffects have a local substrate: A tilt aftereffect field. *Journal of Vision, 10*(13), 5. doi: 10.1167/10.13.5

Dickinson, J. E., Harman, C., Tan, O., Almeida, R. A., & Badcock, D. R. (2012). Local contextual interactions can result in global shape misperception. *Journal of Vision, 12*(11). doi: 10.1167/12.11.3

Dobbins, A., Zucker, S. W., & Cynader, M. S. (1987). Endstopped neurons in the visual cortex as a substrate for calculating curvature. *Nature, 329*(6138), 438-441. doi: 10.1038/329438a0

Dobbins, A., Zucker, S. W., & Cynader, M. S. (1989). Endstopping and curvature. *Vision Research, 29*(10), 1371-1387. doi: 10.1016/0042-6989(89)90193-4

Dodd, J. V., Krug, K., Cumming, B. G., & Parker, A. J. (2001). Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *The Journal of Neuroscience, 21*(13), 4809-4821.

Dumoulin, S. O., Dakin, S. C., & Hess, R. F. (2008). Sparsely distributed contours dominate extra-striate responses to complex scenes. *NeuroImage, 42*(2), 890-901. doi: 10.1016/j.neuroimage.2008.04.266

Dumoulin, S. O., & Hess, R. F. (2006). Modulation of V1 Activity by Shape: Image-Statistics or Shape-Based Perception? *Journal of Neurophysiology, 95*(6), 3654-3664. doi: 10.1152/jn.01156.2005

Durant, S., & Clifford, C. W. G. (2006). Dynamics of the influence of segmentation cues on orientation perception. *Vision Research, 46*(18), 2934-2940. doi: 10.1016/j.visres.2006.02.027

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Boca Raton, FL: Chapman and Hall/CRC.

El-Shamayleh, Y., Kumbhani, R. D., Dhruv, N. T., & Movshon, J. A. (2013). Visual Response Properties of V1 Neurons Projecting to V2 in Macaque. *Journal of Neuroscience, 33*(42), 16594-16605. doi: 10.1523/JNEUROSCI.2753-13.2013

Elder, J. H., Krupnik, A., & Johnston, L. A. (2003). Contour grouping with prior models. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 25*(6), 661-674. doi: 10.1109/TPAMI.2003.1201818

Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic Organization in Human Visual Cortex and the Spatial Precision of Functional MRI. *Cerebral Cortex, 7*(2), 181-192. doi: 10.1093/cercor/7.2.181

Epshtein, B., Lifshitz, I., & Ullman, S. (2008). Image interpretation by a single bottom-up top-down cycle. *Proceedings of the National Academy of Sciences, 105*(38), 14298-14303. doi: 10.1073/pnas.0800968105

Fang, F., Kersten, D., & Murray, S. O. (2008). Perceptual grouping and inverse fMRI activity patterns in human visual cortex. *Journal of Vision, 8*(7). doi: 10.1167/8.7.2

Federer, F., Ichida, J. M., Jeffs, J., Schiessl, I., McLoughlin, N., & Angelucci, A. (2009). Four Projection Streams from Primate V1 to the Cytochrome Oxidase Stripes of V2. *The Journal of Neuroscience, 29*(49), 15455-15471. doi: 10.1523/JNEUROSCI.1648-09.2009

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex (New York, N.Y.: 1991), 1*(1), 1-47.

Felsen, G., Touryan, J., & Dan, Y. (2005). Contextual modulation of orientation tuning contributes to efficient processing of natural stimuli. *Network: Computation in Neural Systems, 16*, 139-149. doi: 10.1080/09548980500463347

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field". *Vision Research, 33*(2), 173-193. doi: 10.1016/0042-6989(93)90156-Q

Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical Surface-Based Analysis: II: Inflation, Flattening, and a Surface-Based Coordinate System. *NeuroImage, 9*(2), 195-207. doi: 10.1006/nimg.1998.0396

Fitzpatrick, D. (2000). Seeing beyond the receptive field in primary visual cortex. *Current opinion in neurobiology, 10*(4), 438-443.

Freeman, J., Donner, T. H., & Heeger, D. J. (2011). Inter-area correlations in the ventral visual pathway reflect feature integration. *Journal of Vision, 11*(4), 15-15. doi: 10.1167/11.4.15

Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nature Neuroscience, 14*(9), 1195-1201. doi: 10.1038/nn.2889

Friston, K. J. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*(1456), 815-836. doi: 10.1098/rstb.2005.1622

Friston, K. J. (2008). Hierarchical Models in the Brain. *PLoS Comput Biol, 4*(11). doi: 10.1371/journal.pcbi.1000211

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and Modulatory Interactions in Neuroimaging. *NeuroImage, 6*(3), 218-229. doi: 10.1006/nimg.1997.0291

Gardelle, V. d., Kouider, S., & Sackur, J. (2010). An oblique illusion modulated by visibility: Non-monotonic sensory integration in orientation processing. *Journal of Vision, 10*(10). doi: 10.1167/10.10.6

Gattass, R., Gross, C. G., & Sandell, J. H. (1981). Visual topography of V2 in the macaque. *The Journal of Comparative Neurology, 201*(4), 519-539. doi: 10.1002/cne.902010405

Gattass, R., Sousa, A. P., & Gross, C. G. (1988). Visuotopic organization and extent of V3 and V4 of the macaque. *The Journal of Neuroscience, 8*(6), 1831-1845.

Geisler, W. S. (2008). Visual Perception and the Statistical Properties of Natural Scenes. *Annual Review of Psychology, 59*(1), 167-192. doi: 10.1146/annurev.psych.58.110405.085632

Geisler, W. S., & Perry, J. S. (2009). Contour Statistics in Natural Images. *Visual Neuroscience, 26*(1), 109-121. doi: 10.1017/S0952523808080875

Geisler, W. S., Perry, J. S., Super, B. J., & Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research, 41*(6), 711-724. doi: 10.1016/S0042-6989(00)00277-7

Georgeson, M. A. (1973). Spatial frequency selectivity of a visual tilt illusion. *Nature, 245*(5419), 43-45. doi: 10.1038/245043a0

Georgeson, M. A., & Harris, M. G. (1984). Spatial selectivity of contrast adaptation: Models and data. *Vision Research, 24*(7), 729-741. doi: 10.1016/0042-6989(84)90214-1

Georgopoulos, A. P., Schwartz, A. B., & Kettner, R. E. (1986). Neuronal Population Coding of Movement Direction. *Science*, 1416-1419.

Gheorghiu, E., & Kingdom, F. A. A. (2006). Luminance-contrast properties of contour-shape processing revealed through the shape-frequency after-effect. *Vision Research, 46*(21), 3603-3615. doi: 10.1016/j.visres.2006.04.021

Gheorghiu, E., & Kingdom, F. A. A. (2007a). Chromatic tuning of contour-shape mechanisms revealed through the shape-frequency and shape-amplitude after-effects. *Vision Research, 47*(14), 1935-1949. doi: 10.1016/j.visres.2007.03.010

Gheorghiu, E., & Kingdom, F. A. A. (2007b). The spatial feature underlying the shape-frequency and shape-amplitude after-effects. *Vision Research, 47*(6), 834-844. doi: 10.1016/j.visres.2006.11.023

Gheorghiu, E., & Kingdom, F. A. A. (2008). Spatial properties of curvature-encoding mechanisms revealed through the shape-frequency and shape-amplitude after-effects. *Vision Research, 48*(9), 1107-1124. doi: 10.1016/j.visres.2008.02.002

Gibson, J. J. (1933). Adaptation, after-effect and contrast in the perception of curved lines. *Journal of Experimental Psychology, 16*(1).

Gibson, J. J., & Radner, M. (1937). Adaptation, after-effect and contrast in the perception of tilted lines. I. Quantitative studies. *Journal of Experimental Psychology, 20*(5).

Gilad, A., Meirovithz, E., & Slovin, H. (2013). Population Responses to Contour Integration: Early Encoding of Discrete Elements and Late Perceptual Grouping. *Neuron, 78*(2), 389-402. doi: 10.1016/j.neuron.2013.02.013

Gilbert, C. D. (1992). Horizontal integration and cortical dynamics. *Neuron, 9*(1), 1-13. doi: 10.1016/0896-6273(92)90215-Y

Gilbert, C. D., Ito, M., Kapadia, M., & Westheimer, G. (2000). Interactions between attention, context and learning in primary visual cortex. *Vision Research, 40*(10–12), 1217-1226. doi: 10.1016/S0042-6989(99)00234-5

Gilbert, C. D., & Wiesel, T. N. (1989). Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *The Journal of Neuroscience, 9*(7), 2432-2442.

Gilbert, C. D., & Wiesel, T. N. (1990). The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. *Vision Research, 30*(11), 1689-1701. doi: 10.1016/0042-6989(90)90153-C

Girard, P., & Bullier, J. (1989). Visual activity in area V2 during reversible inactivation of area 17 in the macaque monkey. *Journal of Neurophysiology, 62*(6), 1287-1302.

Girard, P., Hupé, J. M., & Bullier, J. (2001). Feedforward and Feedback Connections Between Areas V1 and V2 of the Monkey Have Similar Rapid Conduction Velocities. *Journal of Neurophysiology, 85*(3), 1328-1331.

Gitelman, D. R., Penny, W. D., Ashburner, J., & Friston, K. J. (2003). Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. *NeuroImage, 19*(1), 200-207. doi: 10.1016/S1053-8119(03)00058-2

Goddard, E., Clifford, C. W. G., & Solomon, S. G. (2008). Centre-surround effects on perceived orientation in complex images. *Vision Research, 48*(12), 1374-1382. doi: 10.1016/j.visres.2008.02.023

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research, 41*(10–11), 1409-1422. doi: 10.1016/S0042-6989(01)00073-6

Grill-Spector, K., Kushnir, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). Cue-Invariant Activation in Object-Related Areas of the Human Occipital Lobe. *Neuron, 21*(1), 191-202. doi: 10.1016/S0896-6273(00)80526-7

Grill-Spector, K., Kushnir, T., Hendler, T., Edelman, S., Itzchak, Y., Malach, R., & others. (1998). A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Human Brain Mapping, 6*(4), 316-328.

Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology, 35*(1), 96-111.

Grunewald, A., Bradley, D. C., & Andersen, R. A. (2002). Neural Correlates of Structure-from-Motion Perception in Macaque V1 and MT. *The Journal of Neuroscience, 22*(14), 6195-6207.

Hancock, S., & Peirce, J. W. (2008). Selective mechanisms for simple contours revealed by compound adaptation. *Journal of Vision, 8*(7). doi: 10.1167/8.7.11

Händel, B., Lutzenberger, W., Thier, P., & Haarmeier, T. (2007). Opposite Dependencies on Visual Motion Coherence in Human Area MT+ and Early Visual Cortex. *Cerebral Cortex, 17*(7), 1542-1549. doi: 10.1093/cercor/bhl063

Harris, J. P., & Calvert, J. E. (1989). Contrast, spatial frequency and test duration effects on the tilt aftereffect: Implications for underlying mechanisms. *Vision Research, 29*(1), 129-135. doi: 10.1016/0042-6989(89)90179-X

Harrison, L. M., Stephan, K. E., Rees, G., & Friston, K. J. (2007). Extra-classical receptive field effects measured in striate cortex with fMRI. *NeuroImage, 34*(3), 1199-1208. doi: 10.1016/j.neuroimage.2006.10.017

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex. *Science, 293*(5539), 2425-2430. doi: 10.1126/science.1063736

He, D., Kersten, D., & Fang, F. (2012). Opposite Modulation of High- and Low-Level Visual Aftereffects by Perceptual Grouping. *Current Biology, 22*(11), 1040-1045. doi: 10.1016/j.cub.2012.04.026

Heeger, D. J. (1992). Normalization of Cell Responses in Cat Striate Cortex. *Visual Neuroscience, 9*(02), 181-197. doi: 10.1017/S0952523800009640

Herzog, M. H., & Fahle, M. (2002). Effects of grouping in contextual modulation. *Nature, 415*(6870), 433-436. doi: 10.1038/415433a

Herzog, M. H., & Koch, C. (2001). Seeing properties of an invisible object: Feature inheritance and shine-through. *Proceedings of the National Academy of Sciences, 98*(7), 4271-4275. doi: 10.1073/pnas.071047498

Herzog, M. H., Thunell, E., & Ögmen, H. (2015). Putting low-level vision into global context: Why vision cannot be reduced to basic circuits. *Vision Research*. doi: 10.1016/j.visres.2015.09.009

Hess, R. F., & Field, D. (1999). Integration of contours: new insights. *Trends in Cognitive Sciences, 3*(12), 480-486. doi: 10.1016/S1364-6613(99)01410-2
Hildreth, E. C., Ando, H., Andersen, R. A., & Treue, S. (1995). Recovering three-dimensional structure from motion with surface reconstruction. *Vision research, 35*(1), 117-137. doi: 10.1016/0042-6989(94)E0068-V
Huang, P.-C., Chen, C.-C., & Tyler, C. W. (2012). Collinear facilitation over space and depth. *Journal of Vision, 12*(2), 20-20. doi: 10.1167/12.2.20
Huang, P.-C., Hess, R. F., & Dakin, S. C. (2006). Flank facilitation and contour integration: Different sites. *Vision Research, 46*(21), 3699-3706. doi: 10.1016/j.visres.2006.04.025
Hupé, J.-M., James, A. C., Girard, P., & Bullier, J. (2001). Response modulations by static texture surround in area V1 of the macaque monkey do not depend on feedback connections from V2. *Journal of Neurophysiology, 85*(1), 146-163.
Hupé, J.-M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P., & Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature, 394*(6695), 784-787. doi: 10.1038/29537
Ito, H. (2012). Cortical Shape Adaptation Transforms a Circle Into a Hexagon: A Novel Afterimage Illusion. *Psychological Science, 23*(2), 126-132. doi: 10.1177/0956797611422236
Ito, M., & Gilbert, C. D. (1999). Attention Modulates Contextual Influences in the Primary Visual Cortex of Alert Monkeys. *Neuron, 22*(3), 593-604. doi: 10.1016/S0896-6273(00)80713-8
Ito, M., & Komatsu, H. (2004). Representation of Angles Embedded within Contour Stimuli in Area V2 of Macaque Monkeys. *Journal of Neuroscience, 24*(13), 3313-3324. doi: 10.1523/JNEUROSCI.4364-03.2004
Ito, M., Tamura, H., Fujita, I., & Tanaka, K. (1995). Size and position invariance of neuronal responses in monkey inferotemporal cortex. *Journal of Neurophysiology, 73*(1), 218-226.
Ito, M., Westheimer, G., & Gilbert, C. D. (1998). Attention and perceptual learning modulate contextual influences on visual perception. *Neuron, 20*(6), 1191-1197.
Joo, Sung J., Boynton, Geoffrey M., & Murray, Scott O. (2012). Long-Range, Pattern-Dependent Contextual Effects in Early Human Visual Cortex. *Current Biology, 22*(9), 781-786. doi: 10.1016/j.cub.2012.02.067
Joo, Sung J., & Murray, S. O. (2014). Contextual effects in human visual cortex depend on surface structure. *Journal of Neurophysiology, 111*(9), 1783-1791. doi: 10.1152/jn.00671.2013
Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron, 15*(4), 843-856. doi: 10.1016/0896-6273(95)90175-2
Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (1999). Dynamics of spatial summation in primary visual cortex of alert monkeys. *Proceedings of the National Academy of Sciences, 96*(21), 12073-12078. doi: 10.1073/pnas.96.21.12073

Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (2000). Spatial Distribution of Contextual Interactions in Primary Visual Cortex and in Visual Perception. *Journal of Neurophysiology, 84*(4), 2048-2062.

Keeble, D. R. T., & Hess, R. F. (1999). Discriminating local continuity in curved figures. *Vision Research, 39*(19), 3287-3299. doi: 10.1016/S0042-6989(99)00021-8

Kennedy, H., & Bullier, J. (1985). A double-labeling investigation of the afferent connectivity to cortical areas V1 and V2 of the macaque monkey. *The Journal of Neuroscience, 5*(10), 2815-2830.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object Perception as Bayesian Inference. *Annual Review of Psychology, 55*(1), 271-304. doi: 10.1146/annurev.psych.55.090902.142005

Kersten, D., & Yuille, A. (2014). Vision: Bayesian Inference and Beyond. In J. S. Werner & L. M. Chalupa (Eds.), *The new visual neurosciences* (pp. 1263-1278). Cambridge, Massachusetts: The MIT Press.

Khuu, S. K., McGraw, P. V., & Badcock, D. R. (2002). Misperceptions in the peripheral representation of curvature. *Perception, 31 ECVP Abstract Supplement*.

Kim, J., & Horwitz, B. (2008). Investigating the neural basis for fMRI-based functional connectivity in a blocked design: application to interregional correlations and psycho-physiological interactions. *Magnetic Resonance Imaging, 26*(5), 583-593. doi: 10.1016/j.mri.2007.10.011

Knierim, J. J., & Essen, D. C. v. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology, 67*(4), 961-980.

Kobatake, E., & Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology, 71*(3), 856-867.

Koenderink, J. J., & Richards, W. (1988). Two-dimensional curvature operators. *Journal of the Optical Society of America A, 5*(7), 1136. doi: 10.1364/JOSAA.5.001136

Köhler, W., & Wallach, H. (1944). Figural After-Effects. An Investigation of Visual Processes. *Proceedings of the American Philosophical Society, 88*(4), 269-357.

Kohn, A. (2007). Visual Adaptation: Physiology, Mechanisms, and Functional Benefits. *Journal of Neurophysiology, 97*(5), 3155-3164. doi: 10.1152/jn.00086.2007

Kohn, A., & Movshon, J. A. (2003). Neuronal Adaptation to Visual Motion in Area MT of the Macaque. *Neuron, 39*(4), 681-691. doi: 10.1016/S0896-6273(03)00438-0

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, 39, 2729-2737.

Kourtzi, Z., & Kanwisher, N. (2000). Cortical Regions Involved in Perceiving Object Shape. *The Journal of Neuroscience, 20*(9), 3310-3318.

Kourtzi, Z., & Kanwisher, N. (2001). Representation of Perceived Object Shape by the Human Lateral Occipital Complex. *Science, 293*(5534), 1506-1509. doi: 10.1126/science.1061133

Kourtzi, Z., Tolias, A. S., Altmann, C. F., Augath, M., & Logothetis, N. K. (2003). Integration of Local Features into Global Shapes: Monkey and Human fMRI Studies. *Neuron, 37*(2), 333-346. doi: 10.1016/S0896-6273(02)01174-1

Kovács, I. (1996). Gestalten of today: early processing of visual contours and surfaces. *Behavioural Brain Research, 82*(1), 1-11. doi: 10.1016/S0166-4328(97)81103-5
Kovács, I., & Julesz, B. (1993). A Closed Curve Is Much More Than an Incomplete One: Effect of Closure in Figure-Ground Segmentation. *Proceedings of the National Academy of Sciences, 90*(16), 7495-7497.
Kovács, I., & Julesz, B. (1994). Perceptual sensitivity maps within globally defined visual shapes. *Nature, 370*(6491), 644-646. doi: 10.1038/370644a0
Kriegeskorte, N., Sorger, B., Naumer, M., Schwarzbach, J., Van Den Boogert, E., Hussy, W., & Goebel, R. (2003). Human cortical object recognition from a visual motion flowfield. *The Journal of Neuroscience, 23*(4), 1451-1463.
Lamme, V. A. F. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *The Journal of Neuroscience, 15*(2), 1605-1615.
Larsson, J., & Heeger, D. J. (2006). Two Retinotopic Visual Areas in Human Lateral Occipital Cortex. *The Journal of Neuroscience, 26*(51), 13128-13142. doi: 10.1523/JNEUROSCI.1657-06.2006
Lee, J. H., Durand, R., Gradinaru, V., Zhang, F., Goshen, I., Kim, D.-S., . . . Deisseroth, K. (2010). Global and local fMRI signals driven by neurons defined optogenetically by type and wiring. *Nature, 465*(7299), 788-792. doi: 10.1038/nature09108
Lerner, Y., Hendler, T., Ben-Bashat, D., Harel, M., & Malach, R. (2001). A Hierarchical Axis of Object Processing Stages in the Human Visual Cortex. *Cerebral Cortex, 11*(4), 287-297. doi: 10.1093/cercor/11.4.287
Levi, D. M., & Klein, S. A. (2000). Seeing circles: what limits shape perception? *Vision Research, 40*(17), 2329-2339. doi: 10.1016/S0042-6989(00)00092-4
Levitt, J. B., & Lund, J. S. (1997). Contrast dependence of contextual effects in primate visual cortex. *Nature, 387*(6628), 73-76. doi: 10.1038/387073a0
Li, C.-Y., & Li, W. (1994). Extensive integration field beyond the classical receptive field of cat's striate cortical neurons—Classification and tuning properties. *Vision Research, 34*(18), 2337-2355. doi: 10.1016/0042-6989(94)90280-1
Li, W., & Gilbert, C. D. (2002). Global Contour Saliency and Local Colinear Interactions. *Journal of Neurophysiology, 88*(5), 2846-2856. doi: 10.1152/jn.00289.2002
Li, W., Piëch, V., & Gilbert, C. D. (2006). Contour Saliency in Primary Visual Cortex. *Neuron, 50*(6), 951-962. doi: 10.1016/j.neuron.2006.04.035
Li, W., Thier, P., & Wehrhahn, C. (2000). Contextual influence on orientation discrimination of humans and responses of neurons in V1 of alert monkeys. *Journal of Neurophysiology, 83*(2), 941-954.
Li, Z. (1998). A Neural Model of Contour Integration in the Primary Visual Cortex. *Neural Computation, 10*(4), 903-940. doi: 10.1162/089976698300017557
Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences, 6*(1), 9-16. doi: 10.1016/S1364-6613(00)01817-9
Li, Z., & Atick, J. J. (1994). Toward a theory of the striate cortex. *Neural Computation, 6*(1), 127-146.

Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science, 240*(4853), 740-749.
Logothetis, N. K., & Sheinberg, D. L. (1996). Visual Object Recognition. *Annual Review of Neuroscience, 19*(1), 577-621. doi: 10.1146/annurev.ne.19.030196.003045
Lyu, S. (2010). Divisive normalization: Justification and effectiveness as efficient coding transform. *Advances in neural information processing systems, 21*.
Lyu, S. (2011). Dependency reduction with divisive normalization: Justification and effectiveness. *Neural Computation, 23*(11), 2942-2973.
MacKay, D. M. (1956). The Epistemological Problem for Automata. In C. E. Shannon & J. McCarthy (Eds.), *Automata Studies* (pp. 235-251): Princeton University Press.
MacKay, D. M., & Mackay, V. (1974). Do curvature-contingent chromatic aftereffects require "detectors for curvature"? *Vision Research, 14*(11), 1285-1287. doi: 10.1016/0042-6989(74)90230-2
Maffei, L., & Fiorentini, A. (1976). The unresponsive regions of visual cortical receptive fields. *Vision Research, 16*(10), 1131-IN1135. doi: 10.1016/0042-6989(76)90253-4
Magnussen, S., & Greenlee, M. W. (1985). Marathon adaptation to spatial contrast: saturation in sight. *Vision Research, 25*(10), 1409-1411.
Malach, R., Amir, Y., Harel, M., & Grinvald, A. (1993). Relationship between intrinsic connections and functional architecture revealed by optical imaging and in vivo targeted biocytin injections in primate striate cortex. *Proceedings of the National Academy of Sciences, 90*(22), 10469-10473.
Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., . . . Tootell, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences, 92*(18), 8135-8139.
Mannion, D. J., Kersten, D. J., & Olman, C. A. (2014). Regions of Mid-level Human Visual Cortex Sensitive to the Global Coherence of Local Image Patches. *Journal of Cognitive Neuroscience, 26*(8), 1764-1774. doi: 10.1162/jocn_a_00588
Mareschal, I., & Clifford, C. W. G. (2012). Dynamics of unconscious contextual effects in orientation processing. *Proceedings of the National Academy of Sciences, 109*(19), 7553-7558. doi: 10.1073/pnas.1200952109
Mareschal, I., & Clifford, C. W. G. (2013). Spatial structure of contextual modulation. *Journal of Vision, 13*(6), 2-2. doi: 10.1167/13.6.2
Mareschal, I., Morgan, M. J., & Solomon, J. A. (2010). Cortical distance determines whether flankers cause crowding or the tilt illusion. *Journal of Vision, 10*(8). doi: 10.1167/10.8.13
Mareschal, I., Sceniak, M. P., & Shapley, R. M. (2001). Contextual influences on orientation discrimination: binding local and global cues. *Vision Research, 41*(15), 1915-1930.
Markov, N. T., & Kennedy, H. (2013). The importance of being hierarchical. *Current opinion in neurobiology, 23*(2), 187-194. doi: 10.1016/j.conb.2012.12.008
Markov, N. T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., . . . Kennedy, H. (2013). Anatomy of hierarchy: Feedforward and feedback pathways

in macaque visual cortex: Cortical counter-streams. *Journal of Comparative Neurology,* 225-259. doi: 10.1002/cne.23458

Marlin, S. G., Douglas, R. M., & Cynader, M. S. (1991). Position-specific adaptation in simple cell receptive fields of the cat striate cortex. *Journal of Neurophysiology, 66*(5), 1769-1784.

Martin, D., Fowlkes, C., Tal, D., & Malik, J. (2001, 2001). *A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics.* Paper presented at the Eighth IEEE International Conference on Computer Vision, 2001. ICCV 2001. Proceedings.

Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA) - Protein Structure, 405*(2), 442-451. doi: 10.1016/0005-2795(75)90109-9

Maunsell, J. H., & Essen, D. C. V. (1983a). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology, 49*(5), 1127-1147.

Maunsell, J. H., & Essen, D. C. V. (1983b). Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *Journal of Neurophysiology, 49*(5), 1148-1167.

Maunsell, J. H., & Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience, 10*, 363-401. doi: 10.1146/annurev.ne.10.030187.002051

McDonald, J. S., Seymour, K. J., Schira, M. M., Spehar, B., & Clifford, C. W. G. (2009). Orientation-specific contextual modulation of the fMRI BOLD response to luminance and chromatic gratings in human visual cortex. *Vision Research*, 49, 1397-1405.

McGovern, D. P., Roach, N. W., & Webb, B. S. (2014). Characterizing the effects of multidirectional motion adaptation. *Journal of Vision, 14*(13), 2-2. doi: 10.1167/14.13.2

McGuire, B. A., Gilbert, C. D., Rivlin, P. K., & Wiesel, T. N. (1991). Targets of horizontal connections in macaque primary visual cortex. *The Journal of Comparative Neurology, 305*(3), 370-392. doi: 10.1002/cne.903050303

McKeefry, D. J., Watson, J. D. G., Frackowiak, R. S. J., Fong, K., & Zeki, S. (1997). The Activity in Human Areas V1/V2, V3, and V5 during the Perception of Coherent and Incoherent Motion. *NeuroImage, 5*(1), 1-12. doi: 10.1006/nimg.1996.0246

McLaren, D. G., Ries, M. L., Xu, G., & Johnson, S. C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *NeuroImage, 61*(4), 1277-1286. doi: 10.1016/j.neuroimage.2012.03.068

McLeod, P. (1996). Preserved and Impaired Detection of Structure From Motion by a 'Motion-blind" Patient. *Visual Cognition, 3*(4), 363-392. doi: 10.1080/135062896395634

McManus, J. N. J., Li, W., & Gilbert, C. D. (2011). Adaptive shape processing in primary visual cortex. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1105855108

Mendola, J. D., Dale, A. M., Fischl, B., Liu, A. K., & Tootell, R. B. H. (1999). The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *The Journal of Neuroscience, 19*(19), 8560-8572.

Merigan, W. H., Nealey, T. A., & Maunsell, J. H. (1993). Visual effects of lesions of cortical area V2 in macaques. *The Journal of Neuroscience, 13*(7), 3180-3191.

Miyashita, Y., & Chang, H. S. (1988). Neuronal correlate of pictorial short-term memory in the primate temporal cortexYasushi Miyashita. *Nature, 331*(6151), 68-70. doi: 10.1038/331068a0

Moore, C., & Engel, S. A. (2001). Neural Response to Perception of Volume in the Lateral Occipital Complex. *Neuron, 29*(1), 277-286. doi: 10.1016/S0896-6273(01)00197-0

Morgan, M. J., & Baldassi, S. (1997). How the human visual system encodes the orientation of a texture, and why it makes mistakes. *Current Biology, 7*(12), 999-1002.

Morgan, M. J., Mason, A. J. S., & Baldassi, S. (2000). Are there separate first-order and second-order mechanisms for orientation discrimination? *Vision Research, 40*(13), 1751-1763. doi: 10.1016/S0042-6989(00)00015-8

Movshon, J. A., Adelson, E. H., Gizzi, M. S., & Newsome, W. T. (1985). The analysis of moving visual patterns *Pattern recognition mechanisms* (pp. 117-151). New York: Springer.

Movshon, J. A., & Lennie, P. (1979). Pattern-selective adaptation in visual cortical neurones. *Nature, 278*(5707), 850-852. doi: 10.1038/278850a0

Movshon, J. A., & Newsome, W. T. (1996). Visual Response Properties of Striate Cortical Neurons Projecting to Area MT in Macaque Monkeys. *The Journal of Neuroscience*, 16(23), 7733–7741.

Mruczek, R. E. B., Blair, C. D., Strother, L., & Caplovitz, G. P. (2015). The Dynamic Ebbinghaus: motion dynamics greatly enhance the classic contextual size illusion. *Frontiers in Human Neuroscience, 9*. doi: 10.3389/fnhum.2015.00077

Muir, D., & Over, R. (1970). Tilt aftereffects in central and peripheral vision. *Journal of Experimental Psychology, 85*(2), 165-170. doi: 10.1037/h0029509

Muller, J. R., Metha, A. B., Krauskopf, J., & Lennie, P. (2002). Local signals from beyond the receptive fields of striate cortical neurons. *Journal of neurophysiology*, 90(92), 822–831.

Mumford, D. (1992). On the computational architecture of the neocortex: The role of cortico-cortical loops. *Biological Cybernetics, 66*(3), 241-251.

Murray, S. O., Kersten, D., Olshausen, B. A., Schrater, P., & Woods, D. L. (2002). Shape Perception Reduces Activity in Human Primary Visual Cortex. *Proceedings of the National Academy of Sciences, 99*(23), 15164-15169. doi: 10.1073/pnas.192579399

Murray, S. O., Schrater, P., & Kersten, D. (2004). Perceptual grouping and the interactions between visual cortical areas. *Neural Networks, 17*(5-6), 695-705. doi: 10.1016/j.neunet.2004.03.010

Nakayama, K., Shimojo, S., & Ramachandran, V. S. (1990). Transparency: relation to depth, subjective contours, luminance, and neon color spreading. *Perception, 19*(4), 497-513.

Nandy, Anirvan S., Sharpee, Tatyana O., Reynolds, John H., & Mitchell, Jude F. (2013). The Fine Structure of Shape Tuning in Area V4. *Neuron, 78*(6), 1102-1115. doi: 10.1016/j.neuron.2013.04.016

Nelson, S. B. (1991). Temporal interactions in the cat visual system. I. Orientation-selective suppression in the visual cortex. *The Journal of Neuroscience, 11*(2), 344-356.

Newsome, W. T., & Pare, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *The Journal of Neuroscience*, 8(6), 2201–2211.

Nothdurft, H.-C., Gallant, J. L., & Van Essen, D. C. (1999). Response modulation by texture surround in primate area V1: Correlates of "popout" under anesthesia. *Visual Neuroscience, 16*(01), 15-34. doi: null

O'Toole, B., & Wenderoth, P. (1977). The tilt illusion: Repulsion and attraction effects in the oblique meridian. *Vision Research, 17*(3), 367-374. doi: 10.1016/0042-6989(77)90025-6

Olshausen, B. A., & Field, D. J. (1996). Natural image statistics and efficient coding. *Network: Computation in Neural Systems*, 7(2), 333-339.

Orban, G. A. (1997). Visual Processing in Macaque Area MT/V5 and Its Satellites (MSTd and MSTv). In K. S. Rockland, J. H. Kaas & A. Peters (Eds.), *Extrastriate Cortex in Primates* (Vol. 12, pp. 359-434). Boston, MA: Springer US.

Orban, G. A., Sunaert, S., Todd, J. T., Van Hecke, P., & Marchal, G. (1999). Human Cortical Regions Involved in Extracting Depth from Motion. *Neuron, 24*(4), 929-940. doi: 10.1016/S0896-6273(00)81040-5

Over, R., Broerse, J., & Crassini, B. (1972). Orientation illusion and masking in central and peripheral vision. *Journal of Experimental Psychology, 96*(1).

Paradis, A. L., Cornilleau-Pérès, V., Droulez, J., Van De Moortele, P. F., Lobel, E., Berthoz, A., . . . Poline, J. B. (2000). Visual Perception of Motion and 3-D Structure from Motion: An fMRI Study. *Cerebral Cortex, 10*(8), 772-783. doi: 10.1093/cercor/10.8.772

Pasupathy, A., & Connor, C. E. (2002). Population coding of shape in area V4. *Nature Neuroscience, 5*(12), 1332-1338. doi: 10.1038/972

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision, 10*(4), 437-442.

Petrov, Y., & McKee, S. P. (2006). The effect of spatial configuration on surround suppression of contrast sensitivity. *Journal of Vision, 6*(3), 224-238. doi: 10.1167/6.3.4

Pettet, M. W. (1999). Shape and contour detection. *Vision Research, 39*(3), 551-557. doi: 10.1016/S0042-6989(98)00130-8

Pettet, M. W., McKee, S. P., & Grzywacz, N. M. (1998). Constraints on long range interactions mediating contour detection. *Vision Research, 38*(6), 865-879. doi: 10.1016/S0042-6989(97)00238-1

Peuskens, H., Claeys, K. G., Todd, J. T., Norman, J. F., Van Hecke, P., & Orban, G. A. (2004). Attention to 3-D Shape, 3-D Motion, and Texture in 3-D Structure from Motion Displays. *J. Cognitive Neuroscience, 16*(4), 665-682. doi: 10.1162/089892904323057371

Poirier, F. J. A. M., & Wilson, H. R. (2006). A biologically plausible model of human radial frequency perception. *Vision Research, 46*(15), 2443-2455. doi: 10.1016/j.visres.2006.01.026

Polat, U., & Sagi, D. (1993). Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. *Vision Research, 33*(7), 993-999. doi: 10.1016/0042-6989(93)90081-7

Polat, U., & Sagi, D. (1994). The architecture of perceptual spatial interactions. *Vision Research, 34*(1), 73-78. doi: 10.1016/0042-6989(94)90258-5

Press, W. A., Brewer, A. A., Dougherty, R. F., Wade, A. R., & Wandell, B. A. (2001). Visual areas and spatial summation in human visual cortex. *Vision Research, 41*(10), 1321-1332.

Price, N. S. C., & Prescott, D. L. (2012). Adaptation to direction statistics modulates perceptual discrimination. *Journal of Vision, 12*(6), 32-32. doi: 10.1167/12.6.32

Qiu, C., Kersten, D., & Olman, C. A. (2013). Segmentation decreases the magnitude of the tilt illusion. *Journal of Vision, 13*(13), 19-19. doi: 10.1167/13.13.19

R-CoreTeam. (2014). R: A language and environment for statistical computing. *R Foundation for Statistical computing, Vienna, Austria*.

Ramsden, B., Hung, C., & Roe, A. W. (2001). Real and Illusory Contour Processing in Area V1 of the Primate: a Cortical Balancing Act. *Cerebral Cortex, 11*(7), 648-665. doi: 10.1093/cercor/11.7.648

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience, 2*(1), 79-87.

Regan, D., & Beverley, K. I. (1985). Postadaptation orientation discrimination. *Journal of the Optical Society of America A, 2*(2), 147. doi: 10.1364/JOSAA.2.000147

Regan, D., & Hamstra, S. J. (1992). Shape discrimination and the judgement of perfect symmetry: Dissociation of shape from size. *Vision Research, 32*(10), 1845-1864.

Riggs, L. A. (1973). Curvature as a feature of pattern vision. *Science (New York, N.Y.), 181*(4104), 1070-1072.

Rissman, J., Gazzaley, A., & D'Esposito, M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage, 23*(2), 752-763. doi: 10.1016/j.neuroimage.2004.06.035

Roach, N. W., Webb, B. S., & McGraw, P. V. (2008). Adaptation to global structure induces spatially remote distortions of perceived orientation. *Journal of Vision, 8*(3), 31-31. doi: 10.1167/8.3.31

Robol, V., Casco, C., & Dakin, S. C. (2012). The role of crowding in contextual influences on contour integration. *Journal of Vision, 12*(7), 3-3. doi: 10.1167/12.7.3

Robol, V., Tibber, M. S., Anderson, E. J., Bobin, T., Carlin, P., Shergill, S. S., & Dakin, S. C. (2013). Reduced Crowding and Poor Contour Detection in Schizophrenia Are Consistent with Weak Surround Inhibition. *PLoS ONE, 8*(4). doi: 10.1371/journal.pone.0060951

Rockland, K. S., & Lund, J. S. (1983). Intrinsic laminar lattice connections in primate visual cortex. *The Journal of Comparative Neurology, 216*(3), 303-318. doi: 10.1002/cne.902160307

Rockland, K. S., & Pandya, D. N. (1979). Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. *Brain Research, 179*(1), 3-20. doi: 10.1016/0006-8993(79)90485-2

Rockland, K. S., & Virga, A. (1989). Terminal arbors of individual "Feedback" axons projecting from area V2 to V1 in the macaque monkey: A study using immunohistochemistry of anterogradely transported Phaseolus vulgaris-leucoagglutinin. *The Journal of Comparative Neurology, 285*(1), 54-72. doi: 10.1002/cne.902850106

Roe, A. W. (2003). Modular complexity of area V2 in the macaque monkey. *The primate visual system*, 109-138.

Roe, A. W., Parker, A. J., Born, R. T., & DeAngelis, G. C. (2007). Disparity Channels in Early Vision. *Journal of Neuroscience*, 27(44), 11820–11831. doi:11810.11523/JNEUROSCI.14164-11807.12007.

Roelfsema, P. R., Lamme, V. A. F., & Spekreijse, H. (2000). The implementation of visual routines. *Vision Research, 40*(10–12), 1385-1411. doi: 10.1016/S0042-6989(00)00004-3

Roelfsema, P. R., Lamme, V. A. F., & Spekreijse, H. (2004). Synchrony and covariation of firing rates in the primary visual cortex during contour grouping. *Nature Neuroscience, 7*(9), 982-991. doi: 10.1038/nn1304

Saad, Z. S., Reynolds, R. C., Argall, B., Japee, S., & Cox, R. W. (2004, 2004/04//). *SUMA: an interface for surface-based intra- and inter-subject analysis with AFNI.* Paper presented at the IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2004.

Sakai, K., & Hirai, Y. (2002). Neural grouping and geometric effect in the determination of apparent orientation. *JOSA A, 19*(6), 1049-1062.

Sakurai, K. (2014a). Shape distortion illusion of circles without prolonged viewing. *Journal of Vision, 14*(10), 247-247. doi: 10.1167/14.10.247

Sakurai, K. (2014b). Unidirectional shape distortion effect induced by gradation-flash. *Perception, 43 ECVP Abstract Supplement*, 131.

Sakurai, K., & Beaudot, W. (2015). Shape distortion illusion of flashed circles can be induced by dichoptic stimulation. *Journal of Vision, 15*(12), 528. doi: 10.1167/15.12.528

Sandell, J. H., & Schiller, P. H. (1982). Effect of cooling area 18 on striate cortex cells in the squirrel monkey. *Journal of Neurophysiology, 48*(1), 38-48.

Sayim, B., Westheimer, G., & Herzog, M. H. (2010). Gestalt Factors Modulate Basic Spatial Vision. *Psychological Science*. doi: 10.1177/0956797610368811
Sceniak, M. P., Ringach, D. L., Hawken, M. J., & Shapley, R. (1999). Contrast's effect on spatial summation by macaque V1 neurons. *Nature Neuroscience, 2*(8), 733-739. doi: 10.1038/11197
Schallmo, M.-P., Sponheim, S. R., & Olman, C. A. (2013). Abnormal Contextual Modulation of Visual Contour Detection in Patients with Schizophrenia. *PLoS ONE, 8*(6). doi: 10.1371/journal.pone.0068090
Schallmo, M.-P., Sponheim, S. R., & Olman, C. A. (2015). Reduced contextual effects on visual contrast perception in schizophrenia and bipolar affective disorder. *Psychological Medicine, 45*(16), 3527-3537. doi: 10.1017/S0033291715001439
Scharstein, D., & Pal, C. (2007, 2007). *Learning Conditional Random Fields for Stereo.* Paper presented at the IEEE Conference on Computer Vision and Pattern Recognition, 2007. CVPR '07.
Schiller, P. H., & Malpeli, J. G. (1977). The effect of striate cortex cooling on area 18 cells in the monkey. *Brain Research, 126*(2), 366-369. doi: http://dx.doi.org/10.1016/0006-8993(77)90734-X
Schumacher, J. F., & Olman, C. A. (2010). High-resolution BOLD fMRI measurements of local orientation-dependent contextual modulation show a mismatch between predicted V1 output and local BOLD response. *Vision Research, 50*(13), 1214-1224. doi: 10.1016/j.visres.2010.04.005
Schwartz, O., Hsu, A. S., & Dayan, P. (2007). Space and Time in Visual Context. *SSRN eLibrary*.
Schwartz, O., Sejnowski, T., & Dayan, P. (2006). A Bayesian framework for tilt perception and confidence. *Advances in neural information processing systems, 18*.
Schwartz, O., Sejnowski, T. J., & Dayan, P. (2009). Perceptual organization in the tilt illusion. *Journal of Vision, 9*(4). doi: 10.1167/9.4.19
Schwartz, O., & Simoncelli, E. P. (2001). Natural signal statistics and sensory gain control. *Nature Neuroscience, 4*(8), 819-825.
Self, Matthew W., van Kerkoerle, T., Supèr, H., & Roelfsema, Pieter R. (2013). Distinct Roles of the Cortical Layers of Area V1 in Figure-Ground Segregation. *Current Biology, 23*(21), 2121-2129. doi: 10.1016/j.cub.2013.09.013
Sengpiel, F., Sen, A., & Blakemore, C. (1997). Characteristics of surround inhibition in cat area 17. *Experimental Brain Research, 116*(2), 216-228. doi: 10.1007/PL00005751
Sereno, M. E., Trinath, T., Augath, M., & Logothetis, N. K. (2002). Three-Dimensional Shape Representation in Monkey Cortex. *Neuron, 33*(4), 635-652. doi: 10.1016/S0896-6273(02)00598-6
Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., . . . Tootell, R. B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science, 268*(5212), 889-893.
Shapley, R. M., & Xing, D. (2013). Local Circuit Inhibition in the Cerebral Cortex as the source of Gain Control and Untuned Suppression. *Neural networks : the official*

journal of the International Neural Network Society, 37, 172-181. doi: 10.1016/j.neunet.2012.09.005

Shushruth, S., Nurminen, L., Bijanzadeh, M., Ichida, J. M., Vanni, S., & Angelucci, A. (2013). Different Orientation Tuning of Near- and Far-Surround Suppression in Macaque Primary Visual Cortex Mirrors Their Tuning in Human Perception. *The Journal of Neuroscience, 33*(1), 106-119. doi: 10.1523/JNEUROSCI.2518-12.2013

Siegel, R. M., & Andersen, R. A. (1988). Perception of three-dimensional structure from motion in monkey and man. *Nature*, 331, 259-261.

Siegel, R. M., & Andersen, R. A. (1990). The perception of structure from visual motion in monkey and man. *Journal of Cognitive Neuroscience, 2*(4), 306-319.

Sigel, C., & Nachmias, J. (1975). A re-evaluation of curvature-specific chromatic aftereffects. *Vision Research, 15*(7), 829-836. doi: 10.1016/0042-6989(75)90262-X

Sigman, M., Cecchi, G. A., Gilbert, C. D., & Magnasco, M. O. (2001). On a common circle: Natural scenes and Gestalt rules. *Proceedings of the National Academy of Sciences, 98*(4), 1935-1940.

Sillito, A. M., Cudeiro, J., & Jones, H. E. (2006). Always returning: feedback and sensory processing in visual cortex and thalamus. *Trends in Neurosciences, 29*(6), 307-316. doi: 10.1016/j.tins.2006.05.001

Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J., & Davls, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities. *Nature, 378*(6556), 492-496. doi: 10.1038/378492a0

Silverstein, S. M., & Schenkel, L. S. (1997). Schizophrenia as a model of context-deficient cortical computation. *Behavioral and Brain Sciences, 20*(04), 696-697. doi: 10.1017/S0140525X9737160X

Simoncelli, E. P. (2003). Vision and the statistics of the visual environment. *Current opinion in neurobiology, 13*(2), 144-149. doi: 10.1016/S0959-4388(03)00047-3

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience, 24*(1), 1193-1216.

Simoncelli, E. P., & Schwartz, O. (1999). Modeling surround suppression in VI neurons with a statistically-derived normalization model. *Advances in Neural Information Processing Systems 11: Proceedings of the 1998 Confernce, 11*, 153.

Sincich, L. C., Adams, D. L., & Horton, J. C. (2003). Complete flatmounting of the macaque cerebral cortex. *Visual Neuroscience, 20*(06), 663-686.

Sincich, L. C., & Horton, J. C. (2002). Divided by Cytochrome Oxidase: A Map of the Projections from V1 to V2 in Macaques. *Science, 295*(5560), 1734-1737.

Sincich, L. C., & Horton, J. C. (2005). THE CIRCUITRY OF V1 AND V2: Integration of Color, Form, and Motion. *Annual Review of Neuroscience, 28*(1), 303-326. doi: 10.1146/annurev.neuro.28.061604.135731

Smagt, M. J. v. d., Wehrhahn, C., & Albright, T. D. (2005). Contextual Masking of Oriented Lines: Interactions Between Surface Segmentation Cues. *Journal of Neurophysiology, 94*(1), 576-589. doi: 10.1152/jn.00366.2004

Smith, A. T., Greenlee, M. W., Singh, K. D., Kraemer, F. M., & Hennig, J. (1998). The Processing of First- and Second-Order Motion in Human Visual Cortex Assessed by Functional Magnetic Resonance Imaging (fMRI). *The Journal of Neuroscience, 18*(10), 3816-3830.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., . . . Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage, 23 Suppl 1*, S208-219. doi: 10.1016/j.neuroimage.2004.07.051

Snowden, R. J., & Hammett, S. T. (1998). The effects of surround contrast on contrast thresholds, perceived contrast and contrast discrimination. *Vision Research, 38*(13), 1935-1945. doi: 10.1016/S0042-6989(97)00379-9

Snowden, R. J., Treue, S., Erickson, R. G., & Andersen, R. A. (1991). The response of area MT and V1 neurons to transparent motion. *The Journal of Neuroscience, 11*(9), 2768-2785.

Solomon, J. A., Felisberti, F. M., & Morgan, M. J. (2004). Crowding and the tilt illusion: Toward a unified account. *Journal of Vision, 4*(6), 9-9. doi: 10.1167/4.6.9

Solomon, J. A., Watson, A. B., & Morgan, M. J. (1999). Transducer model produces facilitation from opposite-sign flanks. *Vision Research, 39*(5), 987-992. doi: 10.1016/S0042-6989(98)00143-6

Stemmler, M., Usher, M., & Niebur, E. (1995). Lateral interactions in primary visual cortex: a model bridging physiology and psychophysics. *Science, 269*(5232), 1877-1880. doi: 10.1126/science.7569930

Stettler, D. D., Das, A., Bennett, J., & Gilbert, C. D. (2002). Lateral connectivity and contextual interactions in macaque primary visual cortex. *Neuron, 36*(4), 739-750.

Stromeyer, C. F., & Riggs, L. A. (1974). Curvature Detectors in Human Vision? *Science, 184*(4142), 1199-1201. doi: 10.1126/science.184.4142.1199

Strother, L., Killebrew, K. W., & Caplovitz, G. P. (2015). The lemon illusion: seeing curvature where there is none. *Frontiers in Human Neuroscience, 9*. doi: 10.3389/fnhum.2015.00095

Suzuki, S. (2001). Attention-dependent brief adaptation to contour orientation: a high-level aftereffect for convexity? *Vision Research, 41*(28), 3883-3902. doi: 10.1016/S0042-6989(01)00249-8

Suzuki, S. (2003). Attentional selection of overlapped shapes: a study using brief shape aftereffects. *Vision Research, 43*(5), 549-561. doi: 10.1016/S0042-6989(02)00683-1

Suzuki, S., & Cavanagh, P. (1998). A shape-contrast effect for briefly presented stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 24*(5), 1315-1341. doi: 10.1037/0096-1523.24.5.1315

Tadin, D., Kim, J., Doop, M. L., Gibson, C., Lappin, J. S., Blake, R., & Park, S. (2006). Weakened Center-Surround Interactions in Visual Motion Processing in Schizophrenia. *The Journal of Neuroscience, 26*(44), 11403-11412. doi: 10.1523/JNEUROSCI.2592-06.2006

Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience, 19*, 109-139. doi: 10.1146/annurev.ne.19.030196.000545

Tangen, J. M., Murphy, S. C., & Thompson, M. B. (2011). Flashed face distortion effect: Grotesque faces from relative spaces. *Perception, 40*(5), 628-630. doi: 10.1068/p6968

Tanskanen, T., Saarinen, J., Parkkonen, L., & Hari, R. (2008). From local to global: Cortical dynamics of contour integration. *Journal of Vision, 8*(7), 15-15. doi: 10.1167/8.7.15

Tibber, M. S., Bobin, T., Antonova, E., Wright, B., Shergill, S. S., & Dakin, S. C. (2013). Visual surround suppression in schizophrenia. *Frontiers in Psychopathology, 4*. doi: 10.3389/fpsyg.2013.00088

Todd, J. T. (2004). The visual perception of 3D shape. *Trends in Cognitive Sciences, 8*(3).

Tolhurst, D. J., & Thompson, P. G. (1975). Orientation illusions and after-effects: Inhibition between channels. *Vision Research, 15*(8), 967-972.

Tootell, R. B., Mendola, J. D., Hadjikhani, N. K., Ledden, P. J., Liu, A. K., Reppas, J. B., . . . Dale, A. M. (1997). Functional analysis of V3A and related areas in human visual cortex. *The Journal of Neuroscience, 17*(18), 7060-7078.

Tootell, R. B., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., . . . Belliveau, J. W. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *The Journal of neuroscience: the official journal of the Society for Neuroscience, 15*(4), 3215-3230.

Treue, S., Andersen, R. A., Ando, H., & Hildreth, E. C. (1995). Structure-from-motion: Perceptual evidence for surface interpolation. *Vision Research, 35*(1), 139-148. doi: 10.1016/0042-6989(94)E0069-W

Ts'o, D. Y., Gilbert, C. D., & Wiesel, T. N. (1986). Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis. *The Journal of Neuroscience, 6*(4), 1160-1170.

Ullman, S., & Sha'ashua, A. (1988). Structural Saliency: The Detection of Globally Salient Structures Using a Locally Connected Network.

Ungerleider, L. G., & Haxby, J. V. (1994). 'What' and 'where' in the human brain. *Current opinion in neurobiology, 4*(2), 157-165. doi: 10.1016/0959-4388(94)90066-3

Valerio, R., & Navarro, R. (2003). Optimal coding through divisive normalization models of V1 neurons. *Network: Computation in Neural Systems, 14*(3), 579-593.

van der Smagt, M. J., Wehrhahn, C., & Albright, T. D. (2005). contextual masking of oriented lines: interactions between surface segmentation cues. *J. Neurophysiol.*, 576-589.

von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science, 224*(4654), 1260-1262. doi: 10.1126/science.6539501

Wade, A. R., Brewer, A. A., Rieger, J. W., & Wandell, B. A. (2002). Functional measurements of human ventral occipital cortex: retinotopy and colour. *Philosophical Transactions of the Royal Society B: Biological Sciences, 357*(1424), 963-973. doi: 10.1098/rstb.2002.1108

Wallace, G. K. (1969). The critical distance of interaction in the Zöllner illusion. *Perception & Psychophysics, 5*(5), 261-264. doi: 10.3758/BF03209558
Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, 45(44), 205–217. doi:210.1037/h0056880.
Wallisch, P., & Movshon, J. A. (2008). Structure and Function Come Unglued in the Visual Cortex. *Neuron, 60*(2), 195-197. doi: 10.1016/j.neuron.2008.10.008
Wandell, B. A., Brewer, A. A., & Dougherty, R. F. (2005). Visual field map clusters in human cortex. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 360*(1456), 693-707. doi: 10.1098/rstb.2005.1628
Ware, C., & Mitchell, D. E. (1974). The spatial selectivity of the tilt aftereffect. *Vision Research, 14*(8), 735-737. doi: 10.1016/0042-6989(74)90072-8
Wark, B., Lundstrom, B. N., & Fairhall, A. (2007). Sensory adaptation. *Current opinion in neurobiology, 17*(4), 423-429. doi: 10.1016/j.conb.2007.07.001
Watson, A. B. (2014). A formula for human retinal ganglion cell receptive field density as a function of visual field location. *Journal of Vision, 14*(7), 15-15. doi: 10.1167/14.7.15
Watt, R. J., & Andrews, D. P. (1982). Contour curvature analysis: Hyperacuities in the discrimination of detailed shape. *Vision Research, 22*, 449-460.
Webb, B. S. (2005). Early and Late Mechanisms of Surround Suppression in Striate Cortex of Macaque. *Journal of Neuroscience, 25*(50), 11666-11675. doi: 10.1523/JNEUROSCI.3414-05.2005
Webster, M. A. (2011). Adaptation and visual coding. *Journal of Vision, 11*(5), 3-3. doi: 10.1167/11.5.3
Webster, M. A., & Maclin, O. H. (1999). Figural aftereffects in the perception of faces. *Psychonomic Bulletin & Review, 6*(4), 647-653.
Wehrhahn, C., & Dresp, B. (1998). Detection facilitation by collinear stimuli in humans: Dependence on strength and sign of contrast. *Vision Research, 38*(3), 423-428. doi: 10.1016/S0042-6989(97)00141-7
Wehrhahn, C., Li, W., & Westheimer, G. (1996). Patterns That Impair Discrimination of Line Orientation in Human Vision. *Perception, 25*(9), 1053-1064. doi: 10.1068/p251053
Wenderoth, P., & Johnstone, S. (1988). The different mechanisms of the direct and indirect tilt illusions. *Vision Research, 28*(2), 301-312. doi: 10.1016/0042-6989(88)90158-7
Wenderoth, P., & van der Zwan, R. (1989). The effects of exposure duration and surrounding frames on direct and indirect tilt aftereffects and illusions. *Perception & Psychophysics, 46*(4), 338-344.
Westhhmer, G. (1990). Simultaneous orientation contrast for lines in the human fovea. *Vision Research, 30*(11), 1913-1921. doi: 10.1016/0042-6989(90)90167-J
Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling and goodness-of-fit. *Perception and psychophysics*, 1293-1313.
Wilkinson, F., Wilson, H. R., & Habak, C. (1998). Detection and recognition of radial frequency patterns. *Vision Research, 38*(22), 3555-3568. doi: 10.1016/S0042-6989(98)00039-X

Wilson, H. R. (1985). Discrimination of contour curvature: data and theory. *JOSA A, 2*(7), 1191-1199.
Wilson, H. R. (1986). Responses of spatial mechanisms can explain hyperacuity. *Vision Research, 26*(3), 453-469. doi: 10.1016/0042-6989(86)90188-4
Wilson, H. R., McFarlane, D. K., & Phillips, G. C. (1983). Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Research, 23*(9), 873-882. doi: 10.1016/0042-6989(83)90055-X
Wilson, H. R., & Richards, W. A. (1989). Mechanisms of contour curvature discrimination. *JOSA A, 6*(1), 106-115.
Wolfe, J. M. (1984). Short test flashes produce large tilt aftereffects. *Vision Research, 24*(12), 1959-1964. doi: 10.1016/0042-6989(84)90030-0
Woods, R. L., Nugent, A. K., & Peli, E. (2002). Lateral interactions: size does matter. *Vision Research, 42*(6), 733-745. doi: 10.1016/S0042-6989(01)00313-3
Xing, J., & Heeger, D. J. (2000). Center-surround interactions in foveal and peripheral vision. *Vision Research, 40*(22), 3065-3072. doi: 10.1016/S0042-6989(00)00152-8
Xu, H., Dayan, P., Lipkin, R. M., & Qian, N. (2008). Adaptation across the Cortical Hierarchy: Low-Level Curve Adaptation Affects High-Level Facial-Expression Judgments. *The Journal of Neuroscience, 28*(13), 3374-3383. doi: 10.1523/JNEUROSCI.0182-08.2008
Xu, H., Liu, P., Dayan, P., & Qian, N. (2012). Multi-level visual adaptation: Dissociating curvature and facial-expression aftereffects produced by the same adapting stimuli. *Vision Research, 72*, 42-53. doi: 10.1016/j.visres.2012.09.003
Yang, E., Tadin, D., Glasser, D. M., Hong, S. W., Blake, R., & Park, S. (2013). Visual Context Processing in Schizophrenia. *Clinical Psychological Science, 1*(1), 5-15. doi: 10.1177/2167702612464618
Yoon, J. H., Rokem, A. S., Silver, M. A., Minzenberg, M. J., Ursu, S., Ragland, J. D., & Carter, C. S. (2009). Diminished Orientation-Specific Surround Suppression of Visual Processing in Schizophrenia. *Schizophrenia Bulletin*, sbp064. doi: 10.1093/schbul/sbp064
Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: analysis by synthesis? *Trends in Cognitive Sciences, 10*(7), 301-308. doi: 10.1016/j.tics.2006.05.002
Zeki, S., & Shipp, S. (1988). The functional logic of cortical connections. *Nature, 335*(6188), 311-317. doi: 10.1038/335311a0
Zeki, S., Watson, J. D., Lueck, C. J., Friston, K. J., Kennard, C., & Frackowiak, R. S. (1991). A direct demonstration of functional specialization in human visual cortex. *The Journal of Neuroscience, 11*(3), 641-649.
Zhaoping, L. (2014). *Understanding Vision: Theory, Models, and Data*: Oxford University Press.
Zipser, K., Lamme, V. A. F., & Schiller, P. H. (1996). Contextual modulation in primary visual cortex. *The Journal of Neuroscience, 16*(22), 7376-7389.